Humboldt-Universität zu Berlin
Mathematisch-Naturwissenschaftliche Fakultät
Institut für Informatik

Determination and Improvement of Spatial Resolution obtained by Optical Remote Sensing Systems

Dissertation

zur Erlangung des akademischen Grades
Doktors der Ingenieurwissenschaften (Dr.-Ing.)

Präsidentin der Humboldt-Universität zu Berlin

Prof. Dr.-Ing. Dr. Sabine Kunst

Dekan der Mathematisch-Naturwissenschaftlichen Fakultät

Prof. Dr. Elmar Kulke

eingereicht von: Henry Meißner (Dipl.-Ing.)

Gutachter: Prof. Dr. Ralf Reulke
Prof. Dr. Herbert Jahn
Prof. Dr. Heinz-Jürgen Przybilla

eingereicht am: 18.09.2020 verteidigt am: 22.12.2020

for Magdalena

Abstract

There are multiple requirements for objective characterization of resolution parameters for imaging systems and it is beneficial to determine true spatial resolution as a standardized, novel procedure providing all corresponding measurements and confidences.

Although acquisition of resolving power and effective spatial resolution is a well-studied field of research, there are still several scientific questions to be answered when it comes to a standardized (eventually absolute) determination. Therefore, this thesis provides a description of a framework for the imaging process of remote sensing sensors mathematically and methodologically including imaging components and subsequent processes. Furthermore, a detailed review for different structures and methods to measure spatial resolution is included. The aforementioned framework then is utilized to identify related issues to a standardized process obtaining spatial resolution parameters as an image quality criterion to support an upcoming standard DIN 18740-8 "Photogrammetric products – Part 8: Requirements for image quality (quality of optical remote sensing data)".

With respect to define the norm-procedure every measurement influence is quantified where possible and in other cases arrangements are specified to diminish their influence. Moreover, the development of an associated software measurement tool has been accomplished as part of this thesis, which also supports the norm for aerial image quality, spatial resolution in particular.

The framework definition and corresponding tool satisfy the following requirements:

- Reliable and objectively obtained spatial resolution parameters
- Validation and quantification of measurement influences and certainties
- Accurate and verified underlying methodology (enabling self-validation)

As part of a further objective of this thesis, a super-resolution approach to improve spatial resolution of aerial images has been developed and evaluated. The related software tool is able to combine different super-resolution techniques and includes known image quality parameter in subsequent calculations.

The novel super-resolution approach improves spatial resolution of aerial imagery and true ortho-mosaics by taking a set of multiple low-resolved raw images (color filter array), their optimized exterior and interior orientation parameters and the derived 3D-surface as input. Then, one or more super-resolved images are calculated as a hybrid of classic super-resolution method and demosaicing while considering photogrammetric back-projections onto the (data self-derived) 3D-surface. Thereby, limitations of image co-registration with commonly used optical flow approaches can be neglected.

The magnitude of improvement is quantified with the aforementioned standard-

ized measurement method and it is shown that spatial resolution can be improved significantly.

Zusammenfassung

Es existieren zahlreiche Anforderungen, um Auflösungsparameter objektiv zu charakterisieren und es ist lohnend die tatsächliche Auflösung mit Hilfe einer standardisierten Methode zu bestimmen und dabei alle zugehörigen Messunsicherheiten bereitzustellen.

Das Bereitstellen von Parametern bezüglich Auflösungsvermögen und effektiver Auflösung ist ein gut erforschtes Wissenschaftsfeld, dennoch sind noch offene Fragen zu klären, wenn eine standardisierte (schlussendlich absolute) Erhebung angestrebt wird. Zu diesem Zweck ist im Rahmen dieser Arbeit ein Framework definiert und mathematisch und methodologisch beschrieben worden unter Einbeziehung aller untergeordneten Prozesse. Weiterhin liefert sie einen detaillierten Überblick zu den verwendeten Methoden und Strukturen, um räumliche Auflösung zu messen. Das zuvor definierte Framework wird darüber hinaus genutzt, um alle zugehörigen Probleme bezüglich eines genormten Prozesses zu identifizieren und zu lösen. Der so definierte Prozess ist außerdem Teil der bevorstehenden, neuen Norm: DIN 18740-8 "Photogrammetric products – Part 8: Requirements for image quality (quality of optical remote sensing data)".

Im Hinblick auf die definierte Norm sind alle Messeinflüsse an den möglichen Stellen quantifiziert worden und an Stellen, wo dies begründeter maßen nicht möglich ist, wurden Vorkehrungen definiert, die diese Einflüsse mindern. Darüber hinaus wurde ein zugehöriges Softwaretool entwickelt, das ebenfalls die neue Norm unterstützt.

Das erarbeitete Framework und die zugehörige Software erfüllen folgende Spezifikation:

- Zuverlässig und objektiv abgeleitete Auflösungsparameter
- Validierung und Quantifizierung aller Messeinflüsse und Messunsicherheiten
- Exakte und verifizierte, zugrundeliegende Methodologie (selbstvalidierungsfähig)

Als weiterer Schwerpunkt dieser Arbeit wurde ein Verfahren zur Verbesserung der räumlichen Auflösung entwickelt und bewertet. Das zugehörige Softwaretool kombiniert dabei verschiedene Super-Resolution-Ansätze unter Einbeziehung zusätzlicher Kenntnis über die Bildqualität.

Der neuartige Super-Resolution-Ansatz verbessert die räumliche Auflösung von Luftbildern und True-Ortho-Mosaiken indem er ein Set von niedrig aufgelösten Rohbildern (Farbfiltermatrix), deren optimierter, äußerer und innerer Orientierung und die abgeleitete 3D-Oberfläche als Eingangsdaten verwendet. Anschließend werden ein oder mehrere hochaufgelöste Bilder als Kombination von klassischen Super-Resolution-Methoden und De-Mosaikierung berechnet, unter Berücksichtigung der photogrammetrischen Projektionen auf die (dateninhärente) 3D-Oberfläche. Dabei

werden Limitierungen der Bildkoregistrierung mit üblich verwendeten Optical-Flow-Ansätzen überwunden.

Der Grad der Auflösungsverbesserung wird schlussendlich mit der zuvor beschrieben Methode zur Auflösungsbestimmung quantifiziert und es wird gezeigt, dass sich die räumliche Auflösung signifikant verbessert.

Contents

1 Introduction

1.1 Motivation

Remote sensing imagery of the visible spectrum of light is mostly used either for inspection, documentation and classification purposes or to obtain further reconstructions of the observed scene, e.g. digital surface models (DSM), true ortho-photos or ortho-mosaics (TOM) and textured 3D-meshes. All purposes share a mutual requirement, the demand for optimal or best as possible image quality. Parameters describing image quality are necessary for various fields of application (e.g. sensor and mission design, sensor comparison, algorithm development, in-orbit-behaviour of instruments). The effective spatial resolution of the sensor is one important image quality parameter which comprehensively estimates the optical quality of a given sensor-lens combination. Expressing spatial resolution as ground sample distance (GSD), dependent on focal length, pixel-size and flight altitude, is common practice even though the true spatial resolution, ground resolved distance (GRD), may vary tremendously due to numerous effects, e.g. sensor-optic dependent effects or related to external acquisition factors.

Multi-frame aerial imagery obtained for inspection or documentation purposes is wanted to show the finest detail as possible. This also applies to true ortho-mosaics and the texture of 3D-meshes. The 3D-mesh process has been studied thoroughly in photogrammetry and computer vision and can be described, as follows. First, a set of valid interior-orientation (io) and exterior-orientation parameters (eo) have been assigned to every aerial image. Then, these parameters are refined as part of a bundle block adjustment (BBA), usually by optimizing a non-linear least squares problem. Then, 3D-coordinates are reconstructed using dense image matching, e.g. semi-global matching (SGM). Subsequently, the still noisy point cloud is filtered and finally meshed (e.g. Tetrahedral- or Poisson-meshes).

The foremost mentioned bundle block adjustment step and corresponding accuracy verifiably depend on GRD and this allows to conclude that further processes also achieve better or more accurate results using best resolved imagery as input. Furthermore, point cloud filter algorithms are often parameterized with ground sample distance instead of ground resolved distance, usually due to missing knowledge about resolving power to obtain GRD (e.g. scale factor of a point cloud filter-kernel).

From a different perspective, GRD is of prime importance when a minimum resolution has to be guaranteed. An example is the Open Skies treaty. There, the signatories give permission ”for the conduct of observation flights by States Parties over the territories of other States Parties“ obtaining reconnaissance imagery at a predefined minimum resolution that must not be deceeded.

Hence, there are multiple requirements for objective characterization of imaging systems and it is beneficial to determine true spatial resolution as a standardized procedure providing all corresponding measurements and confidences. Although

acquisition of resolving power and effective spatial resolution is a well-studied field of research, there are still several scientific questions to be answered when it comes to a standardized (eventually absolute) determination. This is also research object of a committee of the German Institute for Standardization in preparation of the norm DIN 18740-8 "Photogrammetric products – Part 8: Requirements for image quality (quality of optical remote sensing data)".

Thoroughly executed remote sensing flight campaigns provide highly overlapping aerial images. These campaigns are planned to build a photogrammetric block having side overlaps usually between fifty and sixty percent and up to ninety percent and more along a flight strip, dependent on the aerial camera acquisition frequency. Having highly overlapping imagery draws interest to spatial resolution improvement.

The nature of multi-frame super-resolution is to use several low-resolved images and combine them to one or more super-resolved image. State of the art implementations usually make no or only partial assumptions about image quality and image origination.

Hence, a super-resolution approach using a priori knowledge about the quality of images and all contributing factors should perform better and the spatial resolution improvement can be quantified with the previously defined standardized procedure.

1.2 Objectives

The objective of this thesis is to provide a description of the imaging process of remote sensing sensors mathematically and methodologically including imaging components and subsequent processes. Furthermore, a detailed review for different structures and methods to measure spatial resolution shall be included. The aforementioned description then shall be utilized to identify related issues to a standardized process obtaining spatial resolution parameters as an image quality criteria. With respect to define a standardized procedure every measurement influence shall be quantified where possible and in other cases arrangements shall be specified to diminish their influence.

Moreover, the development of an associated software measurement tool is objective of this thesis, which could be attached to a standard definition of aerial image quality, spatial resolution in particular. The tool should be able to satisfy the following requirements:

- Reliable and objectively obtained spatial resolution parameters
- Validation and quantification of measurement influences and certainties
- Accurate and verified underlying methodology (self-validation)

Further objective of this thesis is the development and evaluation of a super-resolution method to improve spatial resolution of aerial images. The related software tool should be able to combine different super-resolution techniques and include known image

quality parameter in subsequent calculations. Then, the magnitude of improvement shall be quantified with the aforementioned standardized measurement method.

1.3 Main Contributions

The first contribution of this thesis is a reappraisal and an associated software measurement tool for objective characterization of remote sensing imagery regarding spatial resolution as image quality parameter. Therefore a detailed description is delivered, discussing relevant reference structures (bar charts, Slanted-edges, spoke targets) and their methodological challenges, advantages and disadvantages. In addition a novel reference structure, diminishing measurement errors, is conclusively recommended as standard. Thereby, several methods identifying the spoke target's center are featured by utilizing this structure, including a novel (automatic) phase shift approach factoring in test target inclination if required.

A measurement tool, considered to support a standard, must be able to quantify identified standardization aspects (validation) and simultaneously guarantee that underlying methodology is accurate (self-validation). Therefore, the second contribution consists of several validation proofs (e.g. demosaicing method, interpolation methods, motion effects) and self-validation proofs (e.g. comparison to common approaches, model-based verification, structure and exposure-time independence).

Third contribution is a novel super-resolution approach to improve spatial resolution of aerial imagery and true ortho-mosaics. The input is a set of multiple low-resolved raw images (color filter array), their optimized exterior and interior orientation parameters and the derived 3D-surface. Then, one or more super-resolved images are calculated as a hybrid of classic super-resolution method and demosaicing while considering photogrammetric back-projections onto the (data self-derived) 3D-surface. Thereby, limitations of image co-registration with commonly used optical flow approaches can be neglected.

1.4 Outline

The present contribution is structured as follows. Chapter 2 provides a review of mathematical descriptions and criteria for image quality, spatial resolution in particular. Surface reconstruction is a prerequisite of the later described super-resolution approach. Therefore, the 3D reconstruction process is discussed in this chapter, too. This includes bundle block adjustment, dense image matching and point cloud fusion.

In chapter 3 common structures and techniques to obtain spatial resolution measurements are described, compared and discussed regarding their particular advantages. Additionally, a definition of best suited test pattern is presented and explained.

To be able to define a standardized procedure for spatial resolution measurement of remote sensing sensors, all relevant aspects will be investigated, mathematically and methodologically described and quantified in chapter 4 and a conclusive recommen-

dation will be given.

Functional verification of the implemented software measurement tool and its further applications are presented in chapter 5.

The essence of multi-image super-resolution is given in chapter 6 followed by a description of several state of the art methods. Furthermore, a novel approach, combining two independently used techniques, is described and evaluated in this chapter.

The thesis is concluded by summarizing all implications, findings and argumentations and providing an outlook for future work.

2 Related Work

2.1 Image Resolution

2.1.1 Imaging Properties

Image quality of a sensor system is affected by multiple factors. Light beams which are being reflected by an object and detected by a camera sensor partially traverse the atmosphere and lose some of their energy due to diffusion and absorption [1, 2]. In airborne applications (especially UAV) this part can be considered (very) small but gains influence with rising altitude of operation.

Following this, light passes a (complex) lens system where an aperture is integrated and limits the effective solid angles. As a consequence the aperture directly affects the amount of light which in turn determines the amount of photons that reach the sensor plane and contribute to the imaging process. The smaller the aperture is chosen the more diffraction of light limits a sharp optical imaging [3]. On the other hand, if the aperture is chosen too large chromatic aberrations gain influence. Further deviations, caused by the lens system and described as coefficients in the wave aberration function, are distortion, coma, astigmatism, Petzval curvature and spherical aberration [4, 5]. The amount of photons passing through the lens system and reaching the sensor at a distinct time frame directly influences the exposure time needed to create an equivalent sensor signal. In photogrammetry of aerial imagery the exposure time affects a sharp optical imaging in terms of motion blur that is a result of the system's change of location respectively movement whilst the sensor is being exposed. This change of location (translation and/or rotation) can be compensated with stabilized mounts and several aerial camera systems offer some techniques. But most systems (especially for UAV applications) are not equipped with additional parts as they increase total weight limiting flight endurance and operation time. The influence of image blur, comparing imaging under static (laboratory) and dynamic (operating) conditions has been shown in [6].

Another interfering aspect is the gain of shading (or inverse the luminous intensity decrease) starting from the principle point to image corners. This effect is often described as vignetting and is caused by the lens system itself and by the integrated aperture. The vignetting can be measured and corrected as an image processing step while determining Photo Response Non-Uniformity (PRNU) [7].

After the light rays passed the lens system they hit the sensor surface. That is the part of the camera system that creates a digital interpretable signal directly depending on the amount of collected photons during the exposure time window. The quality of that signal is affected by several electronic components (e.g. sensor read-out electronic, analog-digital converter). A measure of this quality is the signal noise ratio (SNR) [8]. The SNR also is characterized by a) the ambient noise level that unavoidably occurs when a semi-conductor is connected to its supply voltage and b) to the photo-effective area of each sensor element (pixel). The larger the effective

area the more photons contribute to the signal assuming identical time frames and therefore increase the signal. The electronic ambient noise can be determined pixel by pixel as part of the Dark Signal Non-Uniformity (DSNU) [7].

During post-processing images of sensors using micro filter patterns (e.g. Bayer-pattern) for color acquisition need to be converted from fractured color array to true color array. The chosen demosaicing (or debayering) method directly affects the effective resolving power [9].

2.1.2 Mathematical Description

Resolving-power investigations date back more than one hundred and forty years. In 1873, Ernst Abbe formulated a resolution limit of optical microscopes in form of minimal distance between elements of a periodic slit pattern [10]. According to his work, in addition to the zero-order maximum of a beam of light, also the first-order maximum has to be captured by the lens-system. The first-order maximum position is directly related to the aperture angle and thus to minimal distance between the slit elements.

In 1874, John W. Strutt (Baron Rayleigh) formulated his resolution limit based on diffraction patterns of two point-like light sources (of similar magnitude) captured by a telescope's circular pupil [11]. Irradiance is expressed as Airy distribution $J_1^2(x)/x^2$ where J_1 is a first order Bessel function and x a cross-section of the sensor plane [12, 13]. Rayleigh showed that, the first minimum in the diffraction pattern occurs at $\theta = 1.22\lambda/D$, where θ is the angular resolution, λ the light source's wavelength and D the aperture diameter. His resolution limit states that, two point-like light sources of similar magnitude are resolved up to a distance where the global maximum of one Airy distribution coincides with the first minimum of the other Airy distribution (see Figure 2.1d). With known distance between source and detector g and using the relation for very small angles $\theta \approx sin\theta = \Delta x/g$, with Δx being the distance between two light sources, the equation can be converted with respect to minimal spatial resolution Δx_{min}:

$$\Delta x_{min} = 1.22\frac{\lambda g}{D} \tag{2.1}$$

In scientific discourse, Rayleigh's resolution limit is often critizised as conservative [15, 16] or even insufficient [17] although it "was not intended as a measure of the actual limit of resolution, but rather as an index of the relative merit of different instruments." [18]. Based on this, the resolution limit "is given, for equal intensities of the two lines [...] by the condition that the central minimum shall just disappear" or in other words "two points of equal brightness can be distinguished if the intensity at the midway point is equal to that at the points" [19]. In mathematical terms the Sparrow criterion for an Airy illuminance distribution I of "two point sources of equal intensity lying off the optical axis a distance β" and substituting 2β with δ_0 can be given as [20]:

$$\frac{\partial^2 I(x, \delta_0)}{\partial x^2} = 0 \tag{2.2}$$

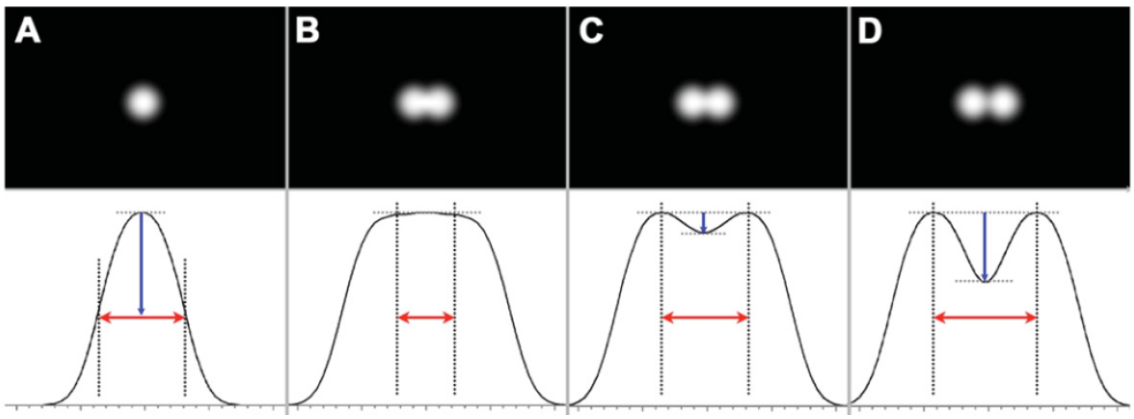


Figure 2.1. (a) FWHM (b) Sparrow (c) Dawes (d) Rayleigh [14]*
*use of this Figure has been approved by and copyright belongs to Dr. Passon

"As the diffraction pattern is symmetric about the origin, all the odd derivatives with respect to x vanish at $x = 0$ [...] which states that the resultant distribution of illuminance undergoes no change in slope." (see Figure 2.1b).

A further resolution limit has been determined empirically with an elaborate and extensive procedure by W.R. Dawes [21]. There, the dip at position intermediate between the two central maxima of the intensity distributions is close to 5% "and is about 18% smaller than that of the Rayleigh limit" [22] (see Figure 2.1c).

A point-like input signal $U(x', y')$ with object space coordinates x' and y' will be spread (or smeared) due to non-ideal imaging properties [8] and creates an output signal $V(x, y)$ with image coordinates x and y:

$$V(x, y) = \iint dx'dy'H(x, y, x', y')U(x', y') \tag{2.3}$$

The spread output signal depends on the system impulse response $H(x, y, x', y')$ which is therefore called point spread function (PSF) [4, 8]. The above-mentioned Airy distribution is one way to define the PSF and another resolution criterion can be described. It is the full width at half the maximum (FWHM) of the point spread function (see Figure 2.1a).

Furthermore, the PSF can also be approximated with a Gaussian function [8]: Then, standard deviation parameter σ of the Gaussian-PSF is an additional criterion to quantify resolving power and is related to FWHM as follows [23, 24, 25]. Starting by assuming a Gaussian-shape function (equation 2.4).

$$H(x) = \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{(x-\mu)^2}{2\cdot\sigma^2}} \tag{2.4}$$

The constant scaling factor $\frac{1}{\sigma\sqrt{2\pi}}$ can be ignored. Applying $H(x) = 0.5$ leads to:

$$e^{-\frac{x_0-\mu}{2\cdot\sigma^2}} = 2^{-1} \tag{2.5}$$

Solving equation (2.5) and assuming function value $H(x_{max})$ occurs at $\mu = 0$ half-maximum points x_0 are found (equation 2.6):

$$x_0 = \pm\sigma\sqrt{2\ln 2} \tag{2.6}$$

The full width at half maximum is then given by:

$$FWHM = x_+ - x_- = 2\sqrt{2\ln 2}\sigma \approx 2.3548\sigma \tag{2.7}$$

Complementary to PSF in image space the optical transfer function $\hat{H}(k_x, k_y)$ with spatial frequencies k_x and k_y specifies the system response in frequency domain and is derived as fourier transfom of the point spread function [8]:

$$\hat{H}(k_x, k_y) = \iint dxdyH(x, y)e^{-j2\pi(k_x\cdot x+k_y\cdot y)} \tag{2.8}$$

As a consequence, the optical transfer function (OTF) is complex valued:

$$\hat{H}(k_x, k_y) = |\hat{H}(k_x, k_y)|e^{-j\Phi_H(k_x,k_y)} \tag{2.9}$$

The amplitude response $|\hat{H}|$ is the modulation transfer function (MTF) and furthermore defined as the quotient of output amplitude and input amplitude. The term $\Phi_H(k_x, k_y)$ is the phase transfer function (PTF) and describes the signal phase shift induced by the imaging system. Implicitly assuming PTF is either constant or otherwise polynomial formulated, often only the modulation transfer function is used to describe resolving power in frequency domain [26, 27, 28, 29, 30]. MTF-value at 50% modulation contrast is used as further resolution criterion and is named effective instantaneous field-of-view (EIFOV) [31]. Assuming a Gaussian-shape function for PSF (equation 2.4) the Fourier-transformed $\tilde{H}(k)$ (MTF with $\tilde{H}(k) = |\hat{H}|$) at fifty percent contrast is formulated as follows (equation 2.10). Again, the constant scaling factor can be ignored.

$$\tilde{H}\,(k) = e^{-2\cdot\pi^2\cdot\sigma^2\cdot k^2} \tag{2.10}$$

With $\tilde{H}(k) = 0.5$ equation (2.10) can be formulated as:

$$2\cdot\pi^2\cdot\sigma^2\cdot k^2 = -\log(0.5) \tag{2.11}$$

Subsequent transposing then gives:

$$k = \sqrt{\frac{-\log(0.5)}{2\cdot\pi^2\cdot\sigma^2}} = \frac{\sqrt{-\log(0.5)/2}}{\pi\cdot\sigma} \tag{2.12}$$

Substituting with C

$$C = \frac{\sqrt{-\log(0.5)/2}}{\pi} \tag{2.13}$$

gives the formula for frequency at fifty percent modulation k_δ (equation 2.14) similar to equation (2.6) and x_0.

$$k \cdot \delta = k_\delta = \frac{C}{\sigma^\delta} \tag{2.14}$$

Finally EIFOV can be calculated with the following equation:

$$EIFOV = \frac{\sigma^\delta}{2C} \approx 2.6667 \cdot \sigma^\delta \tag{2.15}$$

Both criteria, FWHM (equation 2.7) and EIFOV (equation 2.15), are related in their particular domain (image- or frequency-domain) only to parameter σ (PSF respectively MTF) and a similar constant factor.

The value for MTF at 10% modulation contrast often is referred to as resolution limit or cut-off frequency of MTF $\tilde{H}(k) = 0.10$ at spatial frequency k_{MTF10} where it's reciprocal $H(r)$ (PSF) with $r = \sqrt{x^2 + y^2} = \frac{1}{k}$ corresponds to the least resolved scale in image domain. This scale factor multiplied by nominal ground sample distance (GSD) then delivers the least resolved distance and is named ground resolved distance (GRD, see section 3.1.1) [32, 33, 34, 35].

There are two different approaches to formulate this resolution limit mathematically. One is to describe MTF for an aberration-free image and a round entrance pupil as follows [8, 35]:

$$\tilde{H}(k) = \frac{2}{\pi}\left(cos^{-1}\left(\frac{\lambda k}{D}\right) - \frac{\lambda k}{D}\sqrt{\left(1 - \frac{\lambda k}{D}\right)^2}\right) \tag{2.16}$$

By substituting $\frac{1}{r}$ of the Rayleigh resolution limit ($r = 1.22\lambda/D$) for spatial frequency k and solving equation (2.16) the value of MTF at the Rayleigh limit is $\tilde{H}(\frac{1}{r}) \approx 9.0\%$.

A second way to formulate MTF resolution limit is derived from Rayleigh's wavefront criterion (not to be confused with Rayleigh's resolution limit). That wavefront criterion is used for aberrated systems. "The general conclusion is that an aberration between the center and circumference of a quarter period has but little effect upon the intensity at the central point of the image. Translated into mathematical terms, the Rayleigh criterion can be stated in the form

$$max\,|W(x,y)| \leq \frac{\lambda}{4} \tag{2.17}$$

where W(x,y) is the wavefront deviation from a perfect sphere (optical path difference) in the exit pupil and x,y are the dimensionless coordinates in the exit pupil." [36].

It is further shown that, if the wavefront obeys the $\lambda/4$ criterion a Strehl-ratio ≥ 0.80

is being kept. "The Strehl ratio, defined as the ratio of the irradiance at the center of the reference sphere when the aberrations are present to that when they are absent." [37]. MTF is also defined as the relation of image MTF and object MTF [8, 27, 28] and equation (2.18) can be solved for Strehl-ratio > 0.80 :

$$\tilde{H} = \frac{I_{max} - I_{min}}{I_{max} + I_{min}} = \frac{1.00 - 0.81}{1.00 + 0.81} \approx 0.105 \tag{2.18}$$

Therefore, the derived MTF value according to Rayleigh's wavefront criterion is 10.5%.

Even-though, both approaches deliver slightly different MTF values (9% for diffraction limited systems and 10.5% for $\lambda/4$ aberrated systems) it seems common consensus to use the MTF at exactly 10% and it's corresponding spatial frequency k_{MTF10} [28, 30, 32, 33, 38, 39]. The present thesis follows this consensus, utilizing spatial frequency k_{MTF10} and it's reciprocal r_{MTF10} as related distance in image domain. Reliability study and discussion can be found in chapter 4 and 5.

As discussed in section (2.1.1) light beams transverse different physical media from source to detector. Every medium (e.g. atmosphere, lens-system, detector) and process (demosaicing, motion, aliasing) influences the final signal and can separately be expressed as a particular PSF (e.g. H_{lens}, $H_{detector}$, $H_{aliasing}$, etc.) or in frequency domain as a particular MTF (e.g. $\tilde{H}_{lens}$, ,$\tilde{H}_{detector}$, $\tilde{H}_{aliasing}$ etc.). The obtained measured modulation is a combination of all contributing components [8, 27]. In image domain the resulting PSF (H_{all}) is a convolution of image-intensity values (I) and all contributing influences expressed as a particular PSF:

$$H_{all} = I * H_{lens} * H_{detector} * H_{demoisaicing} * H_{aliasing} \cdots \tag{2.19}$$

In frequency domain calculation can be reduced to a product of image spectrum $\tilde{I}$ and all contributing influences expressed as a particular MTF:

$$\tilde{H}_{all} = \tilde{I} \cdot \tilde{H}_{lens} \cdot \tilde{H}_{detector} \cdot \tilde{H}_{demoisaicing} \cdot \tilde{H}_{aliasing} \cdots \tag{2.20}$$

One example, combining two particular modulation transfer functions, is given in Figure (2.2, top). There, the obtained (combined) MTF ($\tilde{H}_{all}$) is the product of diffraction-limited lens-system MTF ($\tilde{H}_{lens}$) and detector MTF ($\tilde{H}_{detector}$). Another example is given in Figure (2.2, bottom). There in contrast, combined MTF ($\tilde{H}_{all}$) is the product of a de-focused lens-system MTF ($\tilde{H}_{lens}$) and detector MTF ($\tilde{H}_{detector}$).

2.2 Bundle Block Adjustment

The processing tool-chain to obtain digital surface models (DSM) from remote sensing data, which are basis of later spatial resolution improvement (see chapter 6), commences by refining parameters of exterior sensor orientation (EO) and/or interior camera orientation (IO). Usually, exterior orientation parameters are given

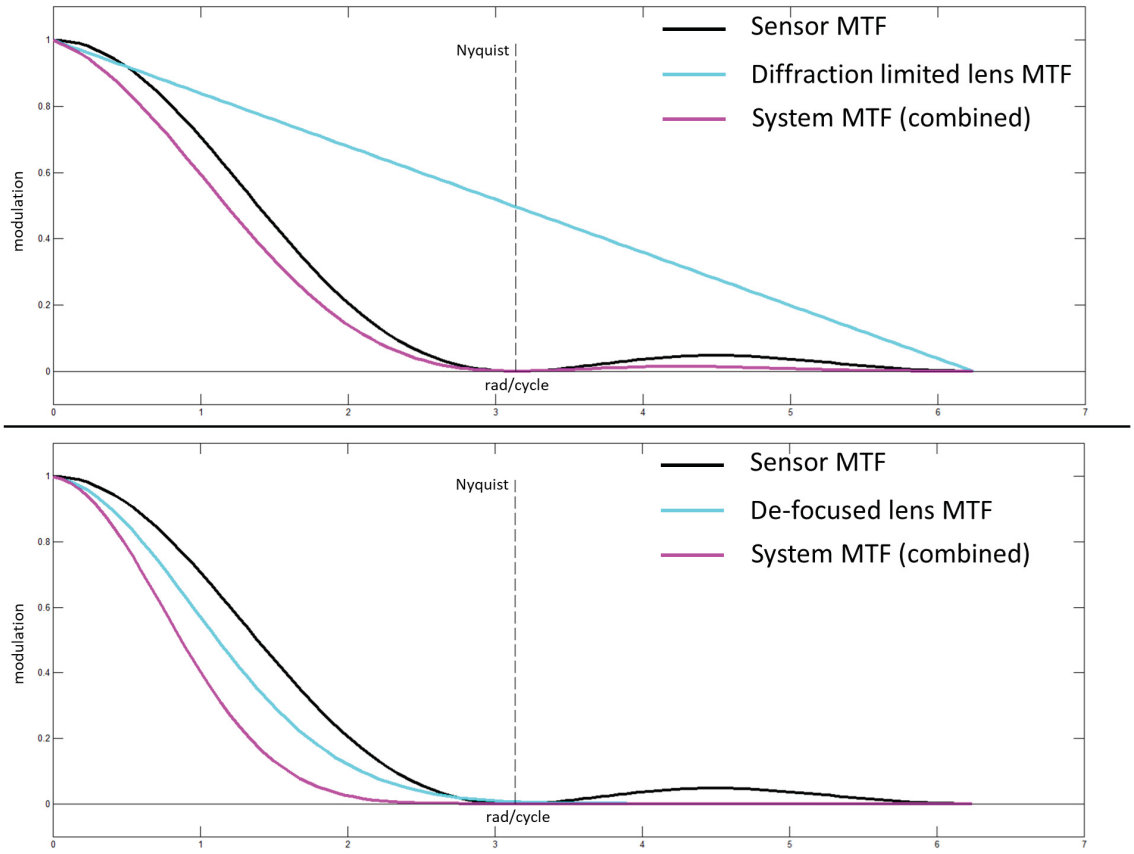


Figure 2.2. Examples for combined system MTFs, (top) Diffraction limited lens MTF, Sensor MTF, combined MTF, (bottom) De-focused lens MTF, Sensor MTF, combined MTF

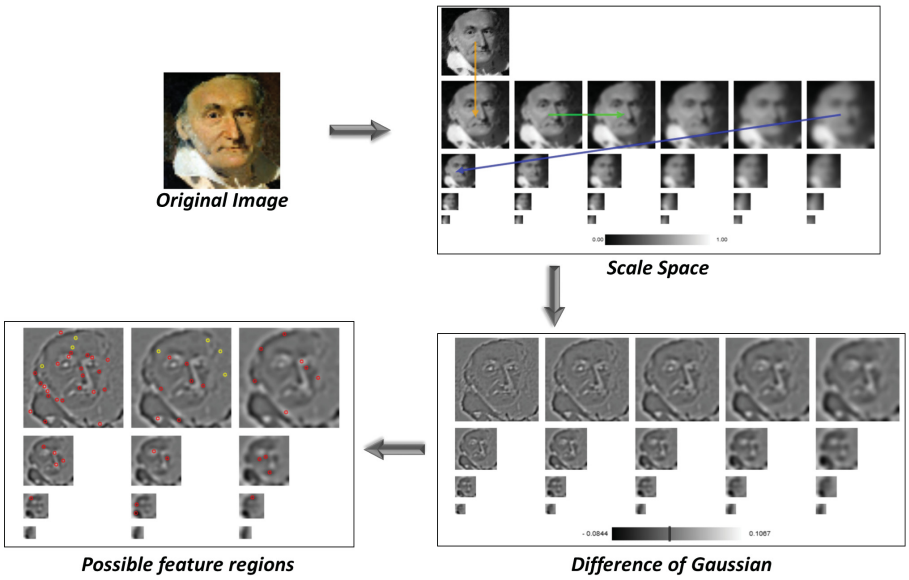


Figure 2.3. SIFT flowchart [41]*, (upper left) original image (upper right) image transformation through scale space (lower right) difference of Gaussian calculation (lower left) identified possible feature regions)
*use of this Figure has been approved by and copyright belongs to Prof. Weitz

having three degrees of freedom (DOF) for translation (X,Y,Z) and three DOF fo rotation (ω,ϕ,κ) Common interior camera parameters are focal length (f) principle point (x_0,y_0) and several distortion paramters (e.g. radially symmetric distortion polynomial with k_0,k_1,k_2,...).

The bundle block adjustment (BBA) approach itself consists of two steps. Finding unique image features, serving as measurements in image space, is the first step (see following section 2.2.1) and minimizing distance between backward projected coarse 3D-estimates and previously determined feature position for homologous features is the second step (see section 2.2.2). The combination of all forward projected rays, projected from image coordinate system to world coordinate system for one unique feature form the shape of a bundle and the process solving a nonlinear least squares problem sets out to deliver a bundle tightened as much as mathematically possible [40] whilst refining IO and EO parameters, hence the name bundle block adjustment.

2.2.1 Feature Extraction

Input for the later applied Levenberg Marquardt minimization (see section 2.2.2) are unique image features provided as result of a robust feature extraction. There are several interest operators delivering reliable features e.g. the Moravec operator [42], Förstner operator [43] and Harris corner detector [44]. However, the most common method to extract features not being prone to scale change has been introduced by Lowe, 1999 [45]. The method is named "Scale Invariant Feature transform" (SIFT)

[46] and the process is described as follows [47]:

"The scale invariant feature transform [...] extracts a set of descriptors from an image. The extracted descriptors are invariant to image translation, rotation and scaling (zoom-out). SIFT descriptors have also proved to be robust to a wide family of image transformations, such as slight changes of viewpoint, noise, blur, contrast changes, scene deformation, while remaining discriminative enough for matching purposes. [...] In order to attain scale invariance, SIFT is built on the Gaussian scalespace, a multiscale image representation simulating the family of all possible zoom-outs through increasingly blurred versions of the input image." (see Figure 2.3, upper right).

"In this popular multiscale framework, the Gaussian convolution acts as an approximation of the optical blur, and the Gaussian kernel approximates the camera's point spread function. Thus, the Gaussian scale-space can be interpreted as a family of images, each of them corresponding to a different zoom factor. [..] Detecting and locating keypoints consists in computing the 3d extrema of a differential operator applied to the scale-space. The differential operator used in the SIFT algorithm is the difference of Gaussians (DoG)." (see Figure 2.3, lower right). The DoG-operator $\sigma\nabla^2 G$ is obtained as a finite difference approximation between two variable scaled Gaussian functions $G(x, y, \sigma)$ and $G(x, y, k\sigma)$ [46]:

$$\sigma\nabla^2 G = \frac{G(x, y, k\sigma) - G(x, y, \sigma)}{k\sigma - \sigma} \tag{2.21}$$

Subsequently 3-dimensional continuous extrema have to be extracted to obtain "a first coarse location of the extrema, which are then refined to subpixel precision using a local quadratic model. [...] Since there are many phenomena that can lead to the detection of unstable keypoints, SIFT incorporates a cascade of tests to discard the less reliable ones. Only those that are precisely located and sufficiently contrasted are retained." (see Figure 2.3, lower left)

The SIFT feature vector v of length $n_{hist} \times n_{hist} \times n_{ori}$ is filled from an array of accumulated histograms h_k^{ij} [47]:

$$v_{((i-1)n_{hist}n_{ori}+(j-1)n_{ori}+k)} = h_k^{ij} \tag{2.22}$$

where $i = 1 \ldots n_{hist}$, $j = 1 \ldots n_{hist}$ and $k = 1 \ldots n_{ori}$. Value n_{hist} is side length of the square-shaped (2D) image histogram and value n_{ori} is the number of reference orientations.

Final step consists of matching a set of found feature vectors from one image to another image [46, 47]. "Let L^A and L^B be the set of descriptors associated to the keypoints detected in images u^A and u^B. The matching is done by considering every descriptor associated to the list L^A and finding one possible match in list L^B. The first descriptor $v^a \in L^A$ is paired to the descriptor $v^b \in L^B$ that minimizes the

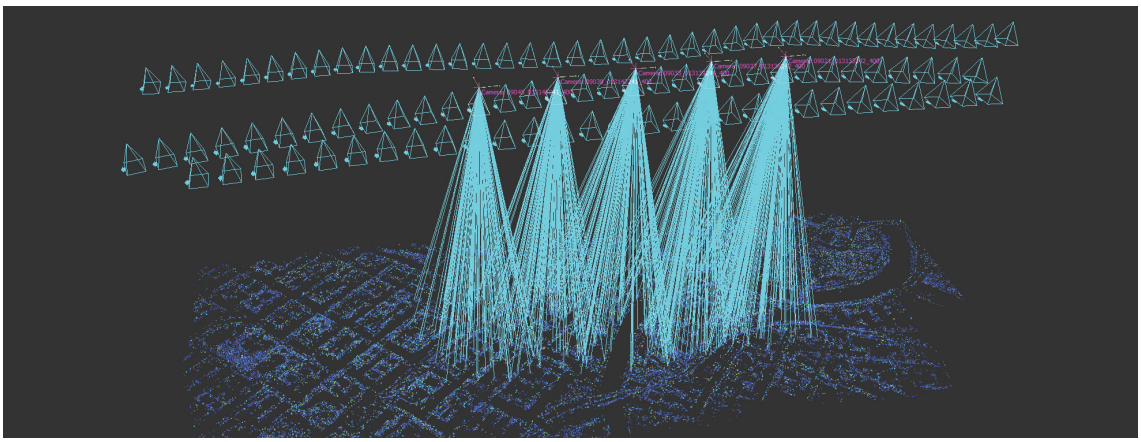


Figure 2.4. Bundle block configuration

Euclidean distance between descriptors“:

$$v^b = \underset{v \in L^B}{\arg\min} ||v - v^a||_2 \tag{2.23}$$

Result is a set of reliable (inter-image linked) scale invariant features that are also resilient to small rotational changes between images and these sets are input for the next step of BBA (see follow section 2.2.2.)

2.2.2 Minimizing Reprojection Error

As already mentioned in the introductory part of this section (sec. 2.2), second step of BBA is to minimize the reprojection error for all found features. Before the 3D world coordinate can be back-projected and minimized a 3D-estimate needs to be calculated. This is done as feature forward projection and can be described as follows. To form a ray, two points are necessary. One point is the feature image location (u, v) and the second point is the principal point (u_0, v_0) at distance equivalent to the focal length f. Subsequently, this ray is translated and rotated in world coordinates according to GNSS/INS-derived EO-parameters. Rays of all images for one distinct feature form the shape of a bundle in more or less close proximity to the real 3D world coordinate depending on EO/IO accuracy (see Figure 2.4). One particular 3D-estimate (using two rays) is the center of the line segment calculated as minimum distance between the two skew rays. As a side not, rays will in almost all cases be skew (and will not intersect) due to noise of EO and/or IO parameters. The final 3D-estimate is calculated as the mean of all possible center estimates for every combination of rays.

Having 3D world coordinates for every feature, the back projection can be calculated. Utilizing a pinhole camera model, projection of an object point in homogeneous coordinates $\mathbf{M} = [X, Y, Z, 1]^T$ to image space in homogeneous coordinates $\mathbf{m} =$

$[u, v, 1]^T$ is obtained as follows [48]:

$$s\,\mathbf{m} \;=\; C\;[R_C^O\;|t_C^O]\;\mathbf{M} \tag{2.24}$$

Rotation matrix $R_C^O = [\mathbf{r_1}\ \mathbf{r_2}\ \mathbf{r_3}]$ and translation vector $t_C^O = [t_x, t_y, t_z]^T$ are the extrinsic parameters. They orientate the camera coordinate systen (C - Camera) to object coordinate system (O - Object). Parameter s is an arbitrary scale factor and camera matrix C, also called intrinsic matrix, is given as:

$$C \;=\; \begin{bmatrix} \alpha & \gamma & u_0 \\ 0 & \beta & v_0 \\ 0 & 0 & 1 \end{bmatrix}\ . \tag{2.25}$$

Parameters α and β reflect the focal length (in pixel) and parameter γ accounts for skew of the image axis (often set to 0). Parameter u_0 and v_0 are coordinates for the principal point. In photogrammetry, principal point corresponds to the optical center of an imaging. Using real lens systems, that point is not necessarily identical to the image center.

Furthermore, real lens systems most often induce radial symmetric distortion causing the imaging of a point to deviate in relation to a pinhole camera model. In practice, radial symmetric distortion is usually modeled using a polynomial of n-th order, where often only the first two terms are used. In that case, the model can be described as:

$$u_r = u + (u - u_0)[k_1(u_n^2 + v_n^2) + k_2(u_n^2 + v_n^2)^2] \tag{2.26}$$

$$v_r = v + (v - v_0)[k_1(u_n^2 + v_n^2) + k_2(u_n^2 + v_n^2)^2] \tag{2.27}$$

True (distorted) position of an image point (u_r, v_r) is the radially shifted position of the calculated coordinates (u, v) depending on principal point (u_0, v_0). The shift distance is determined by distortion parameters k_1 and k_2 and the normalized image coordinates u_n and v_n.

Two equations follow from relations (eq. 2.26) and (eq. 2.27) for every point:

$$\begin{bmatrix} (u - u_0)(u_n^2 + v_n^2) & (u - u_0)(u_n^2 + v_n^2)^2 \\ (v - v_0)(u_n^2 + v_n^2) & (v - v_0)(u_n^2 + v_n^2)^2 \end{bmatrix} \begin{bmatrix} k_1 \\ k_2 \end{bmatrix} = \begin{bmatrix} u_r - u \\ v_r - v \end{bmatrix}, \tag{2.28}$$

The solution can be obtained directly using least squares method for the equation system $Dk = d$, where D is composed of two observations for m points from n images.

Having all features (in image space) and the back-projected image coordinates of the 3D-estimates the final step of BBA "boils down to minimizing the reprojection error between the image locations of observed and predicted image points, which is expressed as the sum of squares of a large number of nonlinear, real-valued functions. Thus, the minimization is achieved using nonlinear least-squares algorithms. Of these, Levenberg–Marquardt has proven to be one of the most successful due to its

ease of implementation and its use of an effective damping strategy that lends it the ability to converge quickly from a wide range of initial guesses. By iteratively linearizing the function to be minimized in the neighbourhood of the current estimate, the Levenberg–Marquardt algorithm involves the solution of linear systems termed the normal equations.“ [49]

The final minimization can be implemented using ”vectors of observations $\underline{z}_i$ predicted by a model $z_i = z_i(x)$,where x is a vector of model parameters. Then nonlinear least squares takes as estimates the parameter values that minimize the weighted Sum of Squared Error (SSE) cost function:” [50]

$$f(x) = \frac{1}{2} \sum_i \Delta z_i(x)^T W_i \Delta z_i(x), \qquad \Delta z_i(x) \equiv \underline{z}_i - z_i(x) \tag{2.29}$$

The technique is ”a common damping (step control) method for nonlinear least squares problems, consisting of adding a multiple λD of some positive definite weight matrix D to the Gauss-Newton Hessian before solving for the step. Levenberg Marquardt uses a simple rescaling based heuristic for setting λ.“ [50].

The algorithm stops either if a minimal, pre-defined step-length ($\lambda \leq \epsilon$) is reached or a number of iterations is exceeded. Then, the minimization results are refined model parameters in form of optimized exterior and/or interior orientation parameters.

On a final note, having a metric camera system (further discussion see section 6.4.1) with reliable long-term stability allows to exclude IO parameters from the nonlinear least squares problem formulation and simultaneously reducing both the number of model parameters and the number of local minima making the minimization more robust.

2.3 Surface Reconstruction

In photogrammetry and computer vision, found and refined image features (see previous section 2.2) are often referred to as a coarse 3D reconstruction of the observed scene. However, it is often necessary to deliver a complete reconstruction containing information as fine and as accurate as possible. Therefore another field of research plays an important role in the 3D reconstruction workflow, namely dense image matching and subsequent point cloud fusion.

2.3.1 Dense Image Matching

Dense image matching aims at finding a corresponding pixel in a match-image for every pixel in a selected base image. The distance between both found pixel locations is called disparity [51]. Obvious limitations (non-reconstructable areas) are occlusions or areas of both images that do not overlap. There are several techniques to produce a dense disparity representation (map) from two images. A quiet simple approach is to use a neighbourhood, surrounding the pixel of interest of the base-frame, and

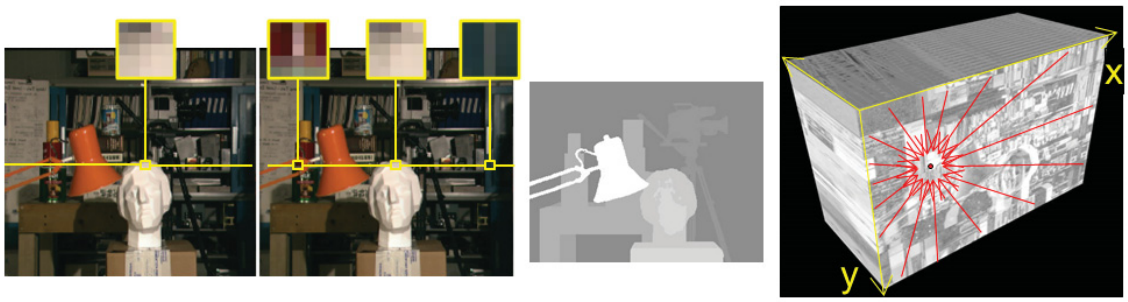


Figure 2.5. Semi Global Matching, (left) rectified base- and match-image with epipolar line and Hamming windows, (center) disparity map, (right) cost cuboid and search-paths, the three cuboid dimension are image coordinates x, y and all possible disparities d

finding the best matching pixel (out of all pixel) in the match-frame. However, this (local) method is prone to blunders especially for images of highly redundant content or repetitive textures. Extending the local approach to a global one delivers more reliable results but increases computation significantly [52].

To account for both problems (runtime and accuracy), Hirschmüller [53] proposed "the Semi-Global Matching (SGM) technique, which offers a good trade off between accuracy and runtime and is therefore well suited for many practical applications".

Starting with rectified images, according to stereo normal case, for every base frame pixel the match-image is traversed alongside the corresponding epipolar line (see Figure 2.5, left). At every epipolar line position the image is scanned at a predefined number of paths (e.g. 8, 16, ...). This step expands the technique to a semi global approach. Furthermore, costs are accumulated for every path, where costs are defined according to pixel similarity (e.g. Hamming distance [54] or mutual information criterion [53]). Final step consists of an energy minimization for all accumulated costs $C(p, D_p)$, where p is the current path and D_p corresponding disparities:

$$E(D) = \sum_{p}(C(p, D_p) + \sum_{q \in N_p} P_1 T\left[|D_p - D_q| = 1\right] + \sum_{q \in N_p} P_2 T\left[|D_p - D_q| > 1\right]) \quad (2.30)$$

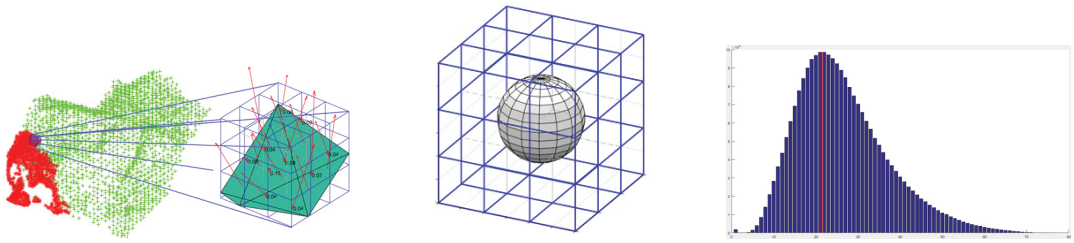
Hirschmüller proposed two smoothness constrains [53]: "The second term adds a constant penalty P_1 for all pixels in the neighborhood [...], for which the disparity changes a little bit (i.e. 1 pixel). The third term adds a larger constant penalty P_2, for all larger disparity changes. Using a lower penalty for small changes permits an adaptation to slanted or curved surfaces. The constant penalty for all larger changes (i.e. independent of their size) preserves discontinuities". The proposed algorithm minimizes the costs for all pixel positions x, y and corresponding disparities d. This setup can be visualized as a cost cuboid (see Figure 2.5, right) and the algorithm finds the path through that cuboid having the lowest overall costs.

Result is a disparity image (or disparity map) for a given base- and match-image combination (see Figure 2.5, center). Subsequently, the disparity images are transformed to world coordinates directly using EO and IO parameters and final result is a single point cloud for every base- and match-image combination.

2.3.2 Point Cloud Fusion

Once the dense matching stage is completed and delivered a combined and (still) noisy point cloud, object space gets spatially sorted by regularly subdividing its axis-aligned bounding box. The length of the edges a of the resulting cubes can be either specified manually in advance if the noise level is known or guessed automatically from a sparse dry-run of the following loop at random. For distortion removal, for each of the voxels the samples inside a $3 \times 3 \times 3$ environment undergo a fast principal component analysis (PCA) to locally estimate the object surface. In order to decide if the current cube is eligible for noise elimination it must be tested efficiently whether it intersects with the approximated planar patch or not. For this purpose a sphere with the radius of half the voxel diagonal $r = \frac{a}{2}\sqrt{3}$ is constructed around its midpoint. When the estimated surface touches the sphere the Euclidean distances of the patch to the points of the cube are calculated and binned to form a histogram. Otherwise the algorithm continues with the next voxel. Since the 26-neighbourhoods of adjacent kernels overlap it is guaranteed by the sampling theorem that the building points are continuously scanned inhibiting structural breaks. Using a sphere to determine whether the locally estimated surface is to be denoisied by a particular voxel or not introduces a small chance that the cube incorrectly accepts the planar patch even though it does not intersect with the element at all. However, since speed is crucial for any volume-based algorithm and the plane normal for the spherical collision test is already available from PCA as the Eigenvector with the smallest Eigenvalue of the underlying covariance matrix the accurate but computationally more expensive intersection test against any of the twelve voxel edges is abandoned in favor of the former.

Having the distance histograms for each voxel noise elimination consists in removing those points from the cube whose distances exceed the value of the class with the maximum sample count. This maximum will be unique if the noise is normally distributed and hence its causes, for instance bundle adjustment issues, do not expose a systematic error. In this case the histogram will follow a discrete version of the χ^2 distribution even though the square root has been taken on the involved differences. Because the square root operation however is a monotonous function the basic profile of the histogram will remain intact. Figure 2.6 illustrates the voxel-based noise removal.



(a) (b) (c)

Figure 2.6. Voxel-based noise removal (a) Surface approximation using PCA (b) neighbourhood and sphere around the kernel for the fast cell intersection test (c) Sample point-to-surface distance histogram

3 Structures and Techniques

Ground resolved distance (GRD) or true ground sample distance (tGSD) is an essential parameter of imaging systems [30, 55], as it defines the detail of information in any image taken by remote sensing sensors. The effective geometric resolution significantly affects photogrammetric processing tool chains. Tie points, mandatory for forming the block geometry, fully rely on feature points (SIFT, SURF, etc.) and the quality parameters of these points however are significantly correlated to image resolution [56]. This is why resolution determination is of such importance to quantify the potential of a sensor-lens-combination.

Although determination of resolving power is a well-studied field of research, there are still some scientific questions to be answered when it comes to a standardized (eventually absolute) determination. This is also research object of a committee of the "German Institute for Standardization" and the given contribution outlines the current state of investigation concerning remote sensing sensors.

Orych [55] provided a description of calibration targets used for high-resolution remote sensing imaging equipment and concluded: "Based on a preliminary analysis, three types of test patterns were selected as possible choices for evaluating the quality of imagery acquired by UAV sensors: bar target, Slanted-edge Test and Siemens-star."

Extending the perspective from UAV-context to a general remote sensing perspective all three approaches must deliver similar or ideally the exact same results for identical images and image regions.

3.1 Bar Test Targets

3.1.1 USAF51 Bar Test Charts

A classic approach is to use defined test targets (e.g. USAF resolution test chart, see Figure 3.1, left) with groups of bars [57]. "The resolving power target used on all tests shall be as follows: The target shall consist of a series of patterns decreasing

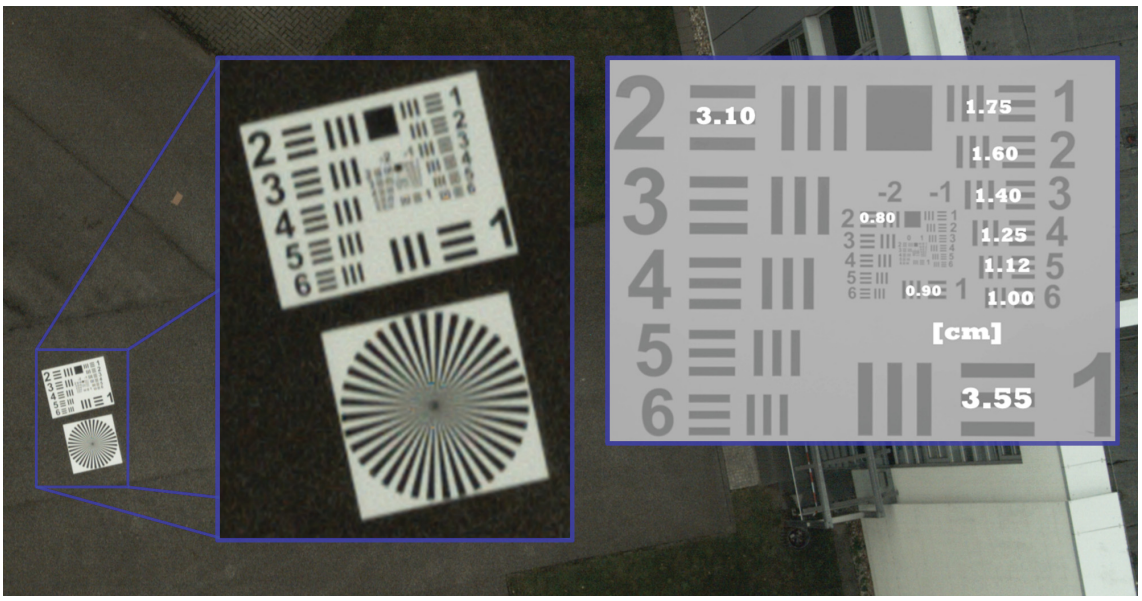


Figure 3.1. Aerial Image of USAF bar test target (left), Corresponding ground resolution [cm] in object space (right)

in size as the $\sqrt{2}$, $\sqrt[3]{2}$, $\sqrt[6]{2}$, with a range sufficient to cover the requirements [...]. The standard target element shall consist of two patterns (two sets of lines) at right angles to each other. Each pattern shall consist of three lines separated by spaces of equal width. Each line shall be five times as long as it is wide." Images of test targets fulfilling these requirements are directly linked to object space metric resolution (see Figure 3.1, right). There, the identified resolution corresponds to the distance between bars of the least discriminable group. The decision whether a group still is discriminable or not strongly relies on viewers' perception. To diminish subjective influence statistically the number of viewers n is chosen to be significant (e.g. $n \geq 10$) and the resulting resolution G_r (GRD, tGSD) is calculated (equation 3.1) as mean of all independent observations G_i.

$$G_r = \frac{1}{n} \sum_{i=1}^{n} G_i \tag{3.1}$$

With knowledge about interior camera parameters (focal length f, pixel-size s) and distance between camera system and test target d the theoretical resolution G_t is calculated as:

$$G_t = \frac{s}{f}\, d \tag{3.2}$$

While G_r is equivalent to GRD or tGSD, the quotient G_q according to equation (3.3) provides another measure for image resolving power.

$$G_q = \frac{G_r}{G_t} \tag{3.3}$$

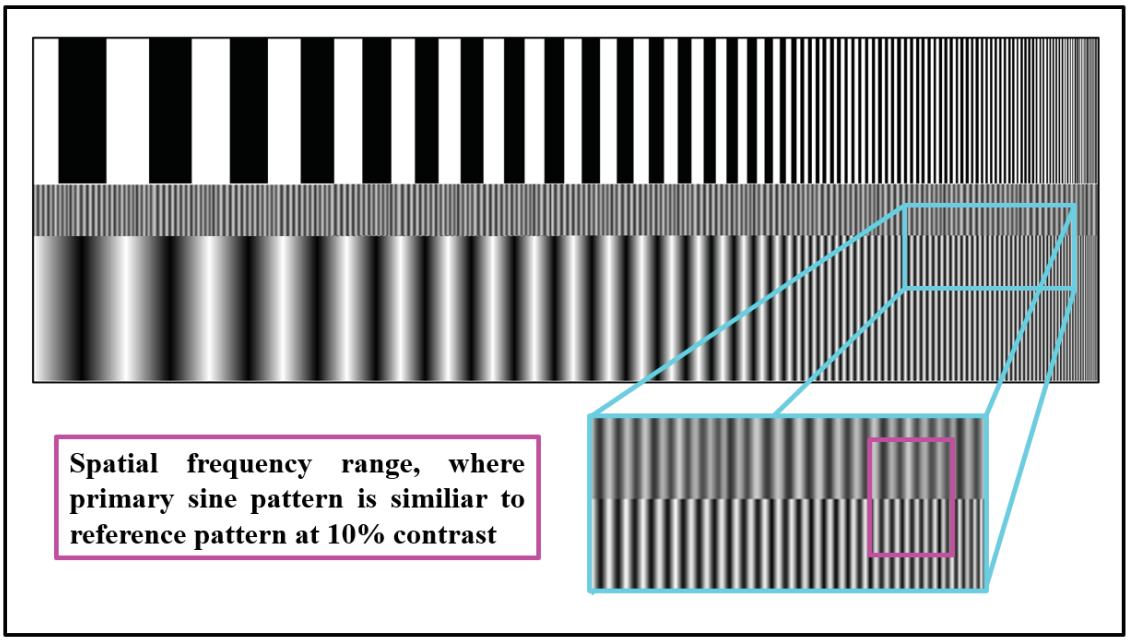


Figure 3.2. Test chart* to determine modulation transfer function in image space, *layout inspired by N. Koren [58]

Usually values for G_q greater than 1 are expected to be calculated. In this case theoretical resolution G_t is better than ultimately determined resolution G_r. Values $G_q \leq 1$ either result due to loss-less transition from object space to image space or indicate image enhancement (e.g. edge-sharpening, color-refinement or super-resolution).

Besides the disadvantage of subjective influence included in this acquisition method values for resolving power are discrete instead of continuous.

3.1.2 Spatial Frequency Charts

The USAF 1951 resolution test charts and its variants are utilized to evaluate resolving power in image-domain (equivalent to time-domain for electromagnetic waves). Other items provide a visual way to evaluate resolving power (in form of MTF-values) directly in frequency-domain.

The Koren test chart [58] for example consists of six bands (see Figure 3.2). "The primary sine pattern [...] and the primary bar pattern [...], are each repeated twice. The other bands are used for reference and visual comparison". Distance between bars of bands 1,3,4 and 6 decrease alongside X-axis or in other words, frequency increases. The magnitude of increase is logarithmically.

To determine spatial frequency for the corresponding MTF, the primary sine pattern (band 1 or 4) has to be compared with the 50% or 10% contrast reference (band 2 or 5). Spatial frequency, where the primary sine pattern is similar to the reference

band (e.g. 10% of contrast), usually is given with an additional confidence interval, since “it is difficult to estimate this frequency with accuracy better than about 15% because MTF rolls off slowly for gaussian blur.”

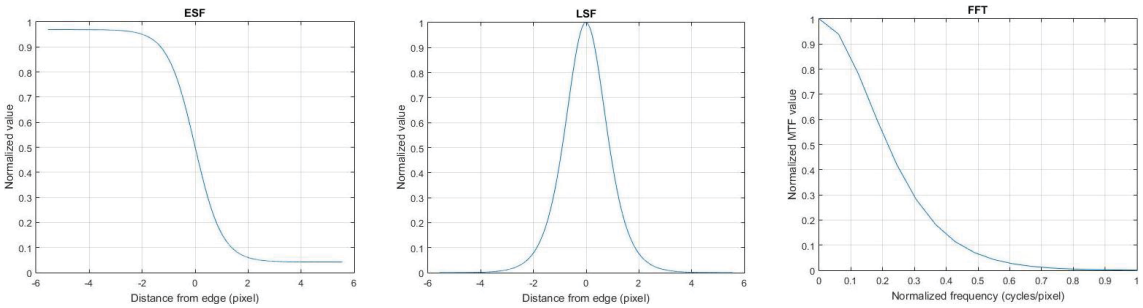
3.2 Slanted-Edge method

The approach to use a slanted-edge to determine spatial resolution parameters is part of the standard ”ISO 12233:2017 Photography — Electronic still picture imaging — Resolution and spatial frequency responses“ [59, 60].

There, the following ten steps are described [61]:

- ”1. Determine a region of interest (ROI) in the image containing a single step edge. (This step, unlike the others, can be performed either manually or automatically.)“
- ”2. Linearize the pixel data by undoing the opto-electronic conversion function (OECF), also known as gamma compression. (Undoing gamma compression is called gamma expansion.)“
- ”3. In the case of a color camera, calculate a weighted sum of the red, green, and blue values to yield a luminance value for each pixel; alternatively, perform the remaining steps separately for each color channel. (Skip this step in the case of a monochrome camera.)“
- ”4. Find the coordinates of points along the intensity step edge.“
- ”5. Fit the parameters of a line to the coordinates.“
- ”6. Project the 2D array of pixel values onto a 1D array known as the edge spread function (ESF).“
- ”7. Differentiate the ESF by convolving with an FIR filter to yield the line spread function (LSF).“
- ”8. Apply a Hamming window function to reduce the effects of noise far from the edge.“
- ”9. Compute the discrete Fourier transform (DFT) of the LSF.“
- ”10. The magnitude of the DFT yields an estimate of the MTF.“

A more detailed description of the presented approach using an edge-step technique can be given as follows [62, 63]. It evaluates the transition between a very homogeneous dark area to a very homogeneous bright area along an extremely sharp, straight edge within the image. The most challenging part of the algorithm is to identify suitable horizontal and vertical edges [63] and to make sure that their position is



(a) Edge Spread Function (b) Line Spread Function (c) Normalized Fourier Transform of the LSF

Figure 3.3. ESF, LSF and Normalized FFT of an Edge.

known to sub-pixel accuracy [62]. Identification of the edges is done automatically either by using a line segment detector [64] or by using a Canny edge detector followed by a Hough transform. Each edge is refined to match the actual transition in the current image as closely as possible, using a custom-built refinement procedure.

After the edges have been located and confirmed to meet the quality standards, their complete profile, spanning their entire length, has to be derived. For each point on the edge, moving along the edge pixel by pixel, the profile following the image's pixel grid is extracted and projected onto the perpendicular to the edge. An alternative approach is to scan and combine multiple perpendicular lines by applying bi-cubic or bi-linear interpolations methods [65].

The thusly obtained projected edge profile is cleaned from blunders, filtered and approximated with a Sigmoid function. The resulting Edge Spread Function (ESF), i.e. the response of the system to this edge [62, 63], is shown in Figure 3.3(a). The numerical derivative of the ESF yields the Line Spread Function (LSF), the response of the system to a line target [62, 63], an example of which is displayed in Figure 3.3(b). Finally, a Fast Fourier Transform (FFT) is applied to the LSF (Figure 3.3(c)) and the normalized magnitude of the result evaluated at the Nyquist frequency (0.5 cycles per pixel) yields the MTF.

3.3 Siemens-Star method

3.3.1 Algorithm

A further method makes use of a priori knowledge of the original scene (well-known Siemens-Star target). First, the mathematical foundation will be presented here followed by a description of the implemented algorithm.

Here, contrast transfer function (CTF), MTF and PSF are approximated by a Gaussian shape function [8, 56] even though an alternative approach is discussed

later (see section 4.3). Coordinate axis X for CTF and MTF is spatial frequency k (equation 3.4) and is calculated as target frequency k_s divided by current scan radius r multiplied by π. Target frequency k_s is constant and equivalent to the number of black-white segments of the well-known Siemens-star.

$$k = \frac{k_s}{\pi r} \tag{3.4}$$

Related (initially discrete) values for contrast transfer function $C_d(k)$ are derived using intensity maxima I_{max} and minima I_{min} for every scanned circle (equation 3.5). Simultaneously the function value is normalized to contrast level C_0 at spatial frequency equal to 0 (infinite radius).

$$C_d(k) = \frac{I_{max}(k) - I_{min}(k)}{I_{max}(k) + I_{min}(k)} * \frac{1}{C_0} \tag{3.5}$$

Continuous function values C are derived by fitting a Gaussian function into discrete input data (equation 3.6).

$$C = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \tag{3.6}$$

According to Coltman [66] the obtained CTF describes the system response to a square wave input while MTF is the system response to a sine wave input. The proposed solution is a normalization with $\frac{\pi}{4}$ followed by series expansion using odd frequency multiples (equation 3.7).

$$\tilde{H}(k) = \frac{\pi}{4}\left[C(f) + \frac{C(3f)}{3} - \frac{C(5f)}{5} + \dots\right] \tag{3.7}$$

MTF describes the effective resolving power in frequency domain while PSF $H(r)$ is the image domain equivalent. For this reason both functions are linked directly by fourier transform (equation 3.8).

$$H(r) \quad \circ\!\!-\!\!\bullet \quad \tilde{H}(k) \tag{3.8}$$

Sequence of the implemented algorithm, that will be official part of the upcoming norm DIN 18740-8 "Photogrammetric products – Part 8: Requirements for image quality (quality of optical remote sensing data)", is illustrated in Figure 3.4 and can be described as follows:

After loading the image, containing a Siemens-star, to the graphical user interface (GUI), the maximum circle-radius and an initial guess of the Siemens-star center position is given by the operator. As a side note, both parameters can be easily adjusted using the mouse in combination with some keyboard-modifier (e.g. CTRL).

Next step consists of finding contrast level C_0 at spatial frequency equal to 0 (infinite radius, equation 3.5). If the proposed structure (see following sub-section 3.3.2)

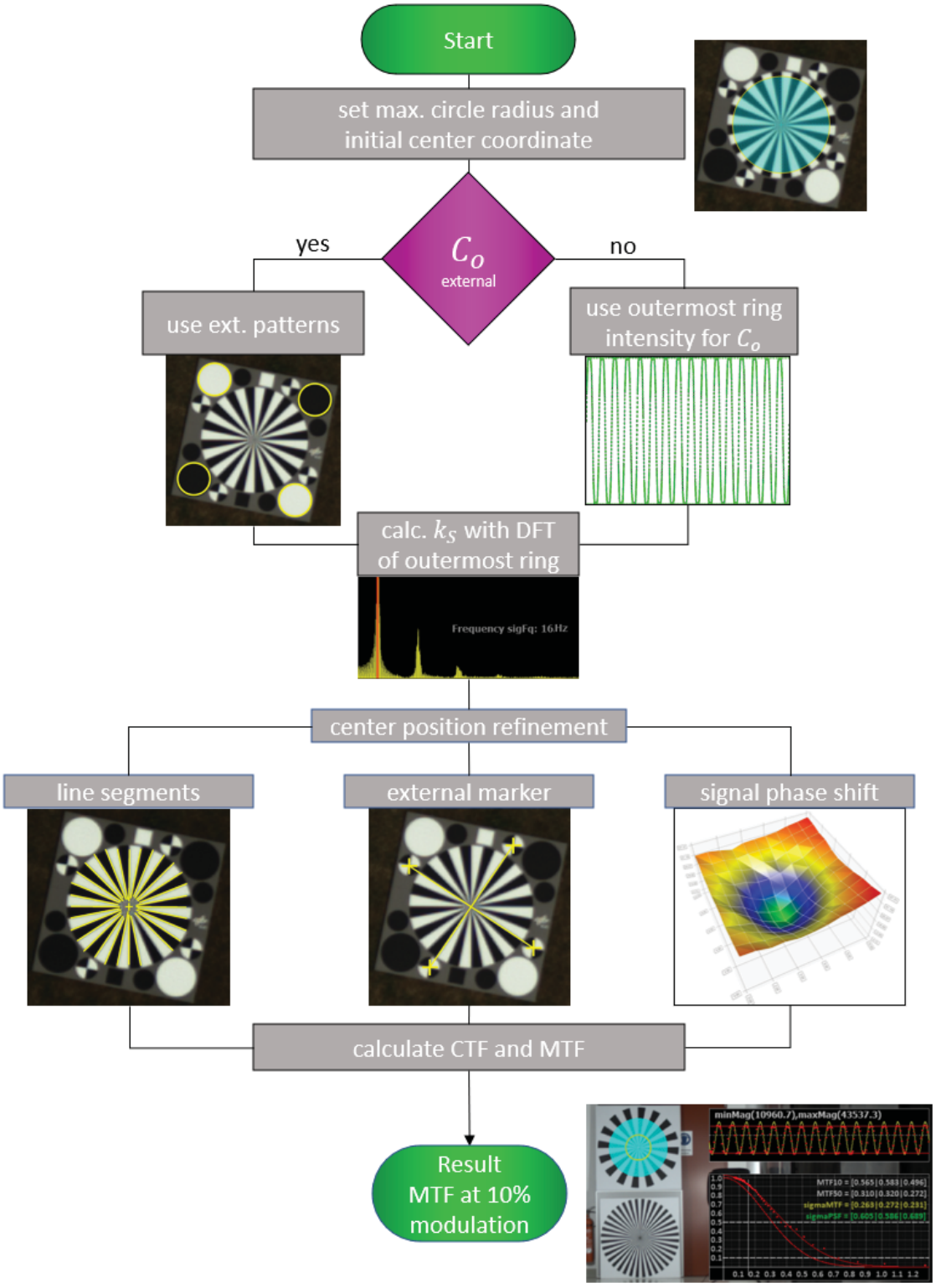


Figure 3.4. MTF determination flowchart

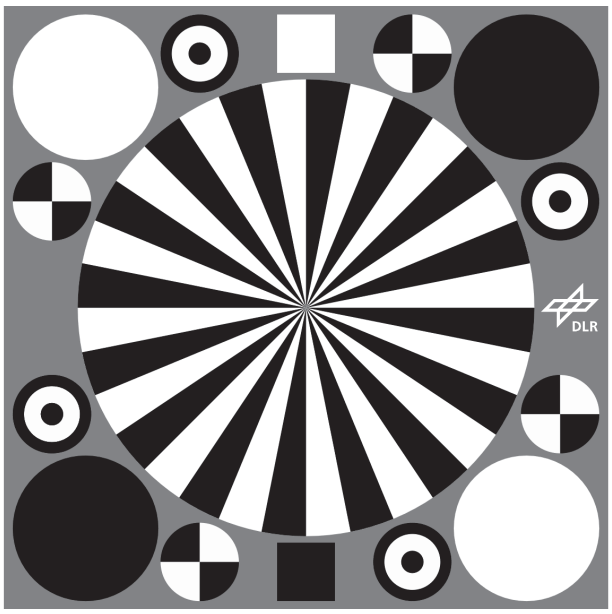


Figure 3.5. Proposed Siemens-star layout

is used, sufficiently large external patterns can be used to obtain C_0, otherwise intensities of the outermost ring are used to calculate C_0.

Determination of target frequency k_s (as part of equation 3.4) is implemented as discrete Fourier transform (DFT) of the outermost ring. Then, k_s is equal to the frequency belonging to largest magnitude.

Next step is center position refinement, see section 4.4 for more details. There will be proven that all three approaches (line segments, external marker and signal phase shift) are interchangeable, delivering very similar results and can be used complementary.

Having the correct center, star frequency k_s and normalization value C_0, MTF is obtained by solving equations 3.5, 3.6 and 3.7 for every ring. Spatial resolution limit is the value of MTF at 10% modulation contrast (see section 2.1.2).

3.3.2 Proposed Structure

Proposed structure of the Siemens-star, supporting objectively obtained measurements for spatial resolution, is constructed containing following features (see Figure 3.5):

- Siemens-star with 32 segments (16 black, 16 white)
- four normalization pattern (2 black and 2 white, redundancy for noise reduction)
- four external markers (rotor shaped, 0.2 px measurement accuracy [67])
- four external markers (circular shaped, for closer ranges - circular fit)

- two external markers (quadratic shaped, one black, one white, to find correct 2D orientation)

The number of star segments is set to 32 (16 black, 16 white). A deeper discussion about different number of segments can be found in section 4.5 where number and overall size can be coupled to intended flight campaign GSD.

Four sufficiently large normalization patterns are included (two black and two white) to obtain values for C_0 (eq. 3.5) more reliable than using the outermost ring. Having two patterns for each intensity increases redundancy and accounts for noise to a certain degree.

Four external markers (rotor shaped) are provided to account for Siemens-star center position refinement (see section 4.4) and the four external markers (circular shaped) are included to account for different approaches to find the marker center.

The two quadratic shaped external markers (one black, one white) have been included in advance for a fully automatic procedure. Identifying both quadratic shaped external markers delivers correct 2D orientation and all other markers may be found automatically using precise knowledge of the designed target.

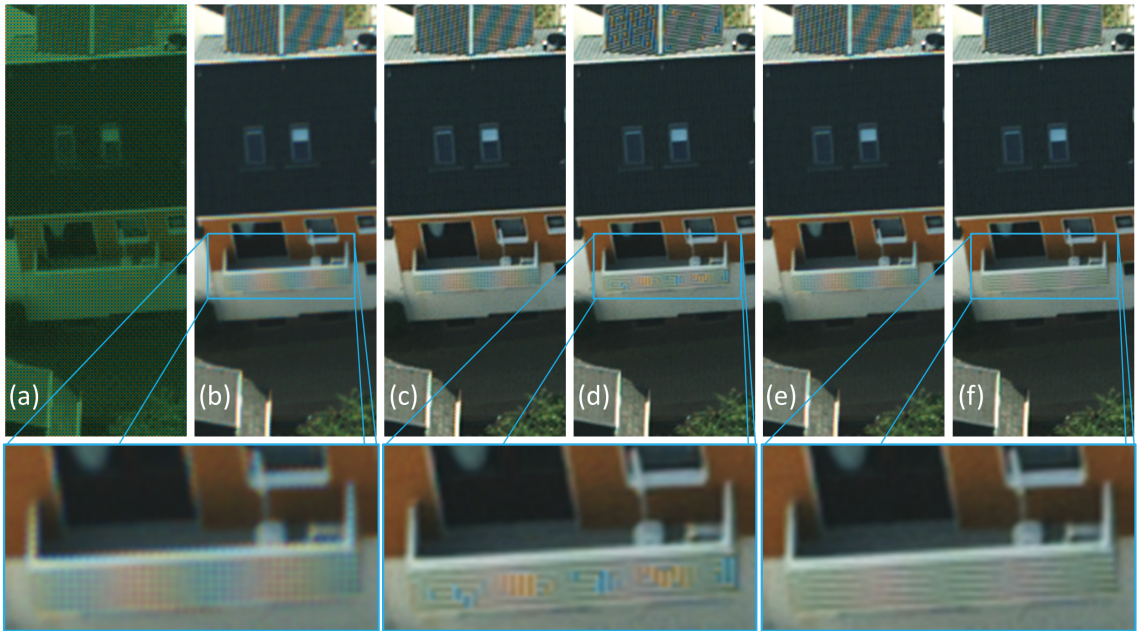


Figure 4.1. Example of different demosaicing algorithms. (a) color filter array (b) bi-linear (c) VNG (d) AHD (e) MHC (f) DCB

4 Standardization Aspects

The importance of image resolution has already been emphasized in 1954 by J. W. Coltman [66]: "The resolving power of an imaging system is a characteristic which often outweighs all others in importance. Unfortunately, there is no general agreement as to the definition of resolving power, and as a consequence, no generally agreed-upon method of measuring it. The usual practice is to image some pattern (such as a series of parallel black and white bars) and reduce the scale of the pattern until the details of the image can no longer be seen. The 'resolving power' is given as a number expressing the scale of the smallest such pattern". Meanwhile, there are several agreed-upon methods, for example the USAF51 bar test chart (see section 3.1.1) as indicated in Coltman's statement and the Slanted-edge method (ISO 12233:2017) in its basic form (see section 3.2). Also, a "sine wave star test chart" is included in this standard. The presented procedure in this thesis (see section 3.3) makes use of square wave test charts but works for sine wave targets, too.

In preparation of DIN 18740-8 and a standardized measurement the following aspects need to be quantified, validated and conclusively discussed.

4.1 Demosaicing Methods

The predominant majority of color cameras use micro filter arrays (one filter evaporated onto every single pixel, so called Bayer-pattern) to capture color information. The array usually is arranged with alternating color filters (e.g. green – red) for a

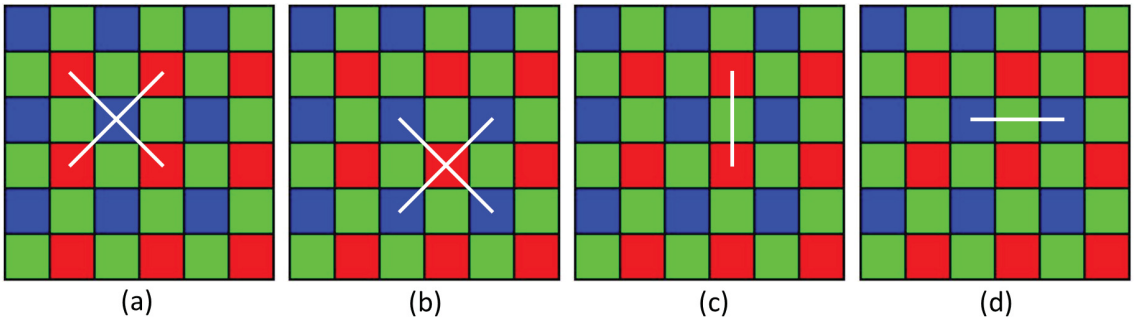


Figure 4.2. Bi-linear demosaicing at different positions (a) red at blue position (b) blue at red position (c) red at green position (d) blue at green position

single sensor line and then a complementary arrangement in the following sensor line (e.g. blue – green), etc. During the color reconstruction process the missing information is determined by interpolating between the neighbourhood values. Hence, demosaicing or debayering is a factor influencing image resolution.

There is a huge variety of demosaicing algorithms out of which several methods have been chosen that are widespread and often used (see Figure 4.1).

4.1.1 Bi-linear Interpolation

The simplest way to restore the missing information is to interpolate each channel separately using neighboring values. Bi-linear interpolation is the most commonly used mode, but it would be possible to use nearest neighbor or bi-cubic interpolation instead. This method is efficient and straight forward to implement, but images will exhibit color fringing at edges. It is worth mentioning that interpolation for red and blue at every green position only is linear instead of bi-linear (see Figure 4.2c and 4.2d). This is one reason why red and blue channel deliver much worse results compared to more sophisticated approaches (see Figure 4.5).

4.1.2 Adaptive Homogeneity-Directed Demosaicing (AHD)

Hirakawa and Parks identify three different classes of artefacts [68]: misguidance color artefacts, interpolation color artefacts and aliasing. They set out to minimize aliasing by using filterbank techniques. Misguidance color artefacts, which arise when the direction of interpolation is erroneously selected (interpolation along an object boundary is preferable to interpolation across the boundary), are addressed through a nonlinear iterative process. The image is interpolated twice - once vertically f_h and once horizontally f_v. The final output f is calculated by combing f_h and f_v based on a homogeneity matrix H_f which aims to minimize color artefacts. AHD can create visually pleasing images, but there are cases, where it gets confused between vertically or horizontally repeating patterns close to the Nyquist frequency (see Figure 4.1, lower mid).

4.1.3 Variable Number of Gradients (VNG)

VNG reduces color fringing by using edge detection [69]. A set of eight gradients is calculated for each pixel by comparing values in the 5 × 5 neighbourhood. The gradient is calculated by summing up the absolute difference of like-colored pixels.

g_1	b_1	g_2	b_2	g_3
r_1	g_4	r_2	g_5	r_3
g_6	b_3	g_7	b_4	g_8
r_4	g_9	r_5	g_{10}	r_6
g_{11}	b_5	g_{12}	b_6	g_{13}

Figure 4.3. Example of gradient calculation for g_7 according to Equation 4.1

The gradient NE (one of eight) at position g_7 is calculated by the following equation:

$$|g_5 - g_9| + |g_3 - g_7| + |b_2 - b_3| + |r_3 - r_5| \tag{4.1}$$

A threshold is used to determine if the pixel lies on a smooth area and averaging can be used to approximate the missing values, or if the pixel lies on a steep gradient, where it is better to use one of two neighboring values.

4.1.4 MHC

MHC is a simpler algorithm than VNG or AHD, it has higher performance than such nonlinear algorithms and doesn't suffer from artefacts due to (sometimes wrong) assumptions about gradients in the image [70, 71]. It works linearly in a 5 × 5 neighbourhood by first filling in values using bilinear interpolation. It tries to analyze local luminance changes by comparing the actual value at the current pixel position to the value arrived at by interpolating same-color neighbors. It then factors a corresponding gain term when calculating the other two color values at the same position (see Figure 4.4).

4.1.5 DCB

Additionally, the iterative DCB approach was chosen to be included in the study [72]. It is largely undocumented, but an open source implementation is available, which performs convincingly [73].

4.1.6 Validation

Several different demosaicing algorithms have been introduced, varying in complexity and performance. The method of choice will depend on processing power, whether results should be visually pleasing or geometrically correct, or other factors such as camera lens design since the resulting PSF affects the correlation between channels

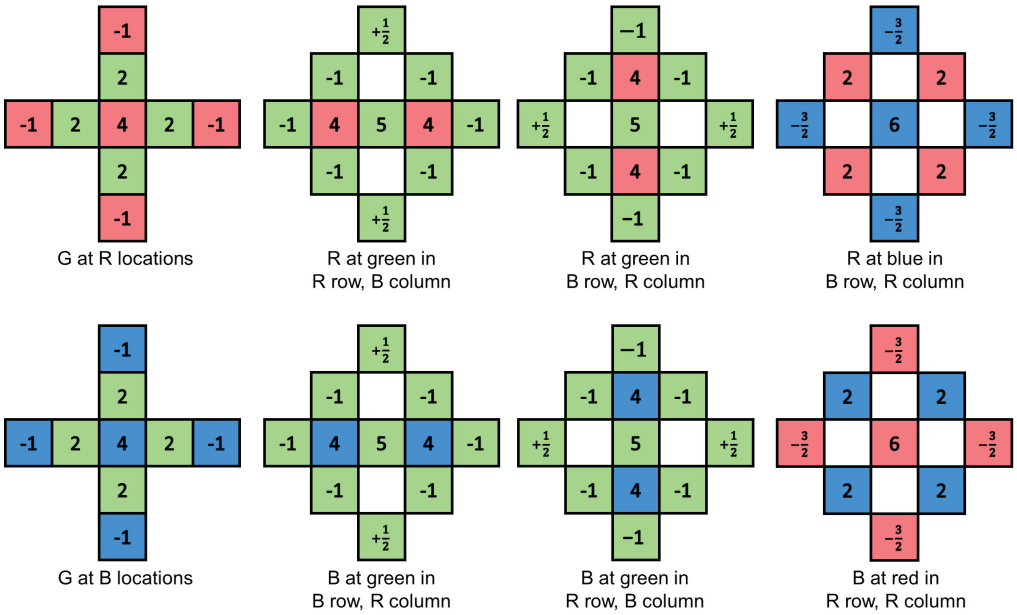


Figure 4.4. Filter coefficients for MHC, different intensity reconstruction patterns for all possible configurations*
*Figure inspired by Malvar, He and Cutler [70]

Figure 4.5. Comparison of different demosaicing methods in terms of image resolution as deviation relatively to VNG approach in percent.

and thus how much information about one channel can be gleaned by analyzing another. A further observation is that poorly calibrated white balance can lead to negative visual artefacts with the edge detecting algorithms for the same reason.

The influence of different demosaicing methods (AHD, DCB, VNG, MHC, Bi-linear) regarding spatial resolution of camera systems suited for remote sensing applications has been investigated for fifty images (see Table 3) of seven different cameras (see Table 1). The images have been acquired under laboratory and operating conditions. The results are given in Figure 4.5 as the deviation relatively to VNG approach in percent with values given in Table 2.

It can be seen that the difference between demosaicing methods and their performance in terms of resulting image resolution partly differ significantly.

The bi-linear approach performs worst, especially for red channel (-26.3%) and blue channel (-27.1%). It shall be mentioned again that bi-linear interpolation for red and blue at every green position is only linear instead of bi-linear (see Figure 4.2c and 4.2d) and this should be a factor of this large inter-channel deviation since the green channel is reconstructed marginally worse compared to VNG (-3.2%). A further conclusion is that every other (more sophisticated) demosaicing approach must have a color harmonization technique implemented and uses a larger neighbourhood to reconstruct color information.

VNG and MHC are almost equal in color reconstruction performance. Results of MHC are slightly better for green channel (green 1.7%) and slightly worse for red ($-0,4\%$) and blue ($-1,2\%$).

The DCB open source implementation performs close to top-rated and in almost all measurements (see Table 3) better than VNG (red 5.2%, green 4.5%, blue 4.9%).

AHD delivers best measurement results (red 6.5%, green 7.2%, blue 5.7%) but, as mentioned above, there are cases, where the AHD approach gets confused between vertically or horizontally repeating patterns close to the Nyquist frequency (see Figure 4.1, lower mid). These reconstruction artefacts are often found as features (tie points) for bundle block adjustment (BBA) and, if not eliminated as outlier, unavoidably falsify the result.

For a standardized procedure it shall be concluded that the used demosaicing method should be mentioned (logged) to enable classification of obtained results. In aerial remote sensing scenarios it is further advised to use AHD approach only if a robust outlier detection during BBA can be guaranteed. The most promising alternative is DCB, even though largely undocumented. If an outlier-free and full documented approach is needed, the choice has to be made between VNG and MHC.

4.2 Signal Scan Interpolation Methods

The presented algorithm in section 3.3 commences by scanning the image in concentric circles and assigning image intensity values to every sampled position. Starting from

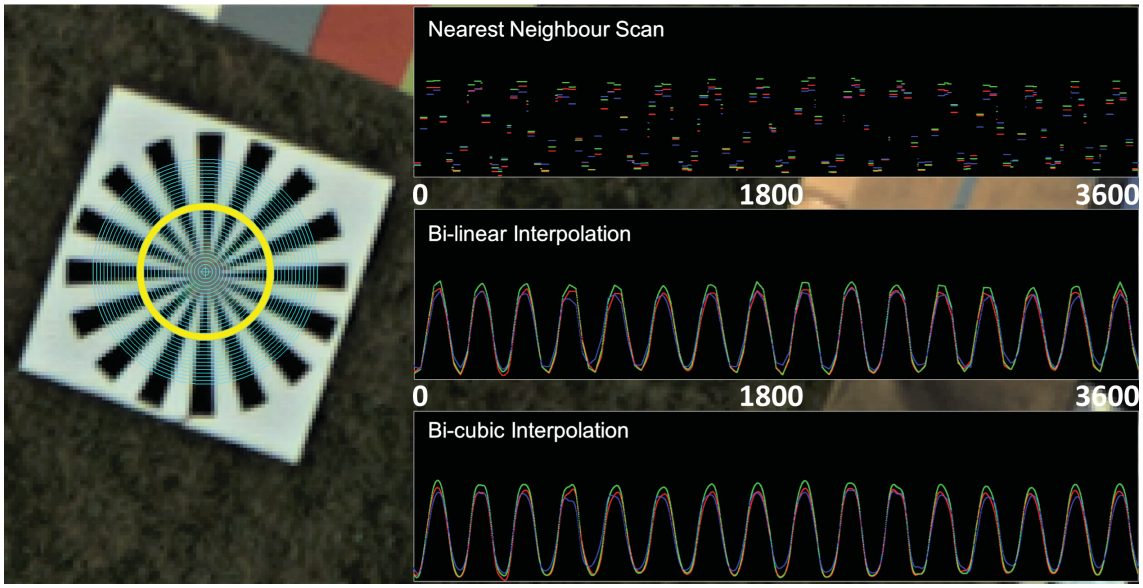


Figure 4.6. Results of different interpolation methods.

correct center location C_{xy} of the star (see section 4.4) the floating point sample position S_{xy} is calculated as follows:

$$S_{xy} = C_{xy} + \begin{pmatrix} x_n \\ y_n \end{pmatrix} \tag{4.2}$$

with

$$\begin{pmatrix} x_n \\ y_n \end{pmatrix} = \begin{pmatrix} cos\alpha_i & -sin\alpha_i \\ sin\alpha_i & cos\alpha_i \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \tag{4.3}$$

Values x and y form the designated start vector depending on radius r (e.g. $x = 0$, $y = r$) and α_i is the rotation angle with predefined sample rate. In this case the increment is set to one tenth of a degree and stored in a vector containing 3600 elements for every scanned circle of the Siemens-star (see Figure 4.6). Each image intensity value at vector position $i \in [0, 3600]$ must be interpolated, assuming image intensity values are defined at pixel center location and the most likely case S_{xy} is different than the pixel center.

In consideration of a standardized approach the underlying interpolation procedure needs to be defined and evaluated. Hereinafter, three different interpolation methods will be described and based on their characteristic a default will be recommended.

4.2.1 Nearest Neighbor

The simplest method to obtain intensity values along scanned circles is the nearest neighbor (or proximal) approach. Neither an averaging nor a weighted function is

being calculated. Instead, every floating point sample position S_{xy} according to equation (4.2) is truncated to integer and the corresponding intensity value at that pixel location is assigned to the related scan vector dependent on radius r (see Figure 4.7a).

Advantage of this approach is that, no calculated artefacts are induced (e.g. signal overshoot).

It is disadvantageous that, image aliasing effects are not smoothed and the scanned output signal is rather discontinuous (see Figure 4.6, top right). These aliasing effects, due to non-consideration of sub-pixel position, produce minima I_{min} and maxima I_{min} (see equation 3.5) that are too low respectively too high.

4.2.2 Bi-linear Interpolation

A more sophisticated way to acquire intensity values is bi-linear interpolation [65]. This method searches for a function of the form:

$$f(x,y) = \sum_{i=0}^{1}\sum_{j=0}^{1} a_{ij}x^i y^j = a_{00} + a_{10}x + a_{01}y + a_{11}xy \tag{4.4}$$

where every internal cell is mapped to the unit square (see Figure 4.7b) and four conditions need to be satisfied $f^A = f(0,0)$, $f^B = f(1,0)$, $f^C = f(1,1)$ and $f^D = f(0,1)$. Inverting the 4×4 system gives:

$$a_{00} = f^A \tag{4.5}$$
$$a_{10} = f^B - f^A \tag{4.6}$$
$$a_{01} = f^D - f^A \tag{4.7}$$
$$a_{11} = f^C + f^A - f^B - f^D \tag{4.8}$$

In contrast to nearest neighbor, the bi-linear interpolation approach smooths aliasing effects (see Figure 4.6, center right). The obtained circle-scan is continuous even across cell boundaries but the derivative is not.

4.2.3 Bi-cubic Interpolation

The lowest order two-dimensional interpolation approach which maintains continuity of the function and its first derivative is bi-cubic interpolation [65]. This method searches for a function of the form:

$$f(x,y) = \sum_{i=0}^{3}\sum_{j=0}^{3} a_{ij}x^i y^j \tag{4.9}$$

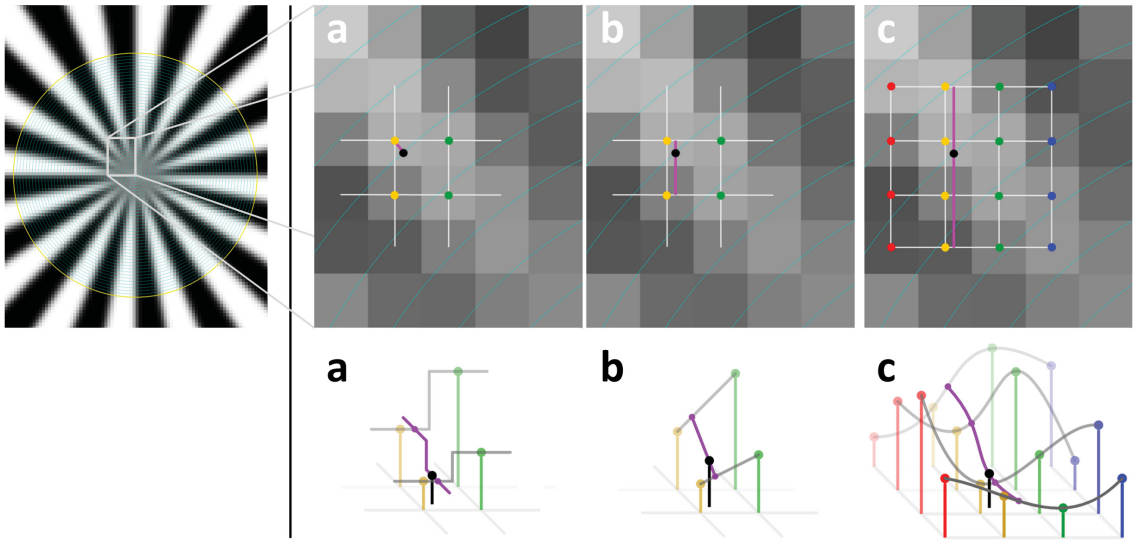


Figure 4.7. Visualization of different interpolation methods. a) Nearest neighbor, b) Bi-linear, c) Bi-Cubic. (bottom row visualization under CC BY-SA 4.0 [74])

and the first derivatives:

$$f_x(x,y) = \sum_{i=0}^{3}\sum_{j=0}^{3} a_{ij} i x^{i-1} y^j \tag{4.10}$$

$$f_y(x,y) = \sum_{i=0}^{3}\sum_{j=0}^{3} a_{ij} x^i j y^{j-1} \tag{4.11}$$

$$f_{x,y}(x,y) = \sum_{i=0}^{3}\sum_{j=0}^{3} a_{ij} i x^{i-1} j y^{j-1} \tag{4.12}$$

where $f_x(x,y)$ and $f_y(x,y)$ are derivatives in x- respectively y-direction and $f_{xy}(x,y)$ the cross-derivative. To solve the interpolation problem sixteen coefficients have to be determined. Function values of the unit square corners themselves, values of the first derivative in each direction and the cross-derivative sum up to sixteen know values and the resulting 16×16 system of equations can be solved.

An efficient way to implement bi-cubic interpolation is to split the problem into cubic interpolation in one dimension (e.g. x-direction) and then the perpendicular direction using the previously obtained cubic interpolated values [75]. The polynomial is constructed as a Catmull-Rom spline [76] where the first derivative at a particular location on the unit square (e.g. $f_x(0,0)$) is calculated as follows:

$$f_x(0,0) = \frac{f(1,0) - f(-1,0)}{2} \tag{4.13}$$

The bi-cubic approach delivers smooth and continuous scans (see Figure 4.6, lower right) but uses a large area to obtain the measurement data (see Figure 4.7c). Also, this method does not prevent overshoots, which means that signal amplitudes between the original measurement can be lower or higher than actual intensity values, especially when sharpening filters have been applied to the image.

4.2.4 Validation

When looking at the signal scans (see Figure 4.6) it can be seen that the nearest neighbor scan is discontinuous due to the unsophisticated method. In contrast, bi-linear interpolation is continuous even across cell boundaries and does not suffer from overshoots. Bi-cubic interpolation delivers smoothest scans and even the derivative is continuous across cell boundaries but at the same time it needs a 4×4 neighbourhood instead of 2×2 neighbourhood (bi-linear) or only one pixel (nearest neighbor) to obtain corresponding intensity values. This large scan extent affects the measurement, especially for signals close to Nyquist-frequency. Then, the extent partly lies in unresolved area and the other part lies in resolved area. The received results are to optimistic since resolved intensities are induced into unresolved ones. This can be seen when looking at measurement data of fifty images taken by seven different cameras (see Table 4). On average, the nearest neighbor approach delivers results 6% (red 6.2%, green 6.2%, blue 5.8%) more optimistic than bi-linear interpolation. The difference between bi-linear and bi-cubic is similar. On average, results are more optimistic by 5% (red 4.8%, green 5.2%, blue 4.8%).

As a compromise between smoothness, no signal overshoot and medium area of influence bi-linear shall be considered as sweet spot and should be the default signal scan method.

4.3 MTF Approximation

Values for MTF and corresponding spatial frequency according to the approach described in section 3.3 (equation 3.7) are discrete, depending on step size between concentric rings (e.g. 1 pixel). The resolution limit is equivalent to spatial frequency k_{MTF10} where MTF is at 10% modulation contrast (see section 2.1.2) and there are several ways to obtain this value.

4.3.1 Gaussian Fit

One way to approximate MTF is to fit a Gaussian function into obtained measurements [8] (see Figure 4.8). Main advantage is that this approach only needs one value to describe the approximated Gaussian function. It is the standard deviation parameter σ since mean μ of the Gaussian function is at $k_0 = \mu = 0$ and value for MTF at spatial frequency $\tilde{H}(k_0) = 1$. The implementation of a Gaussian fit has been done by iteratively solving a non-linear regression problem until either a number of maximum iterations (e.g. $n = 50$) is exceeded or the difference between model

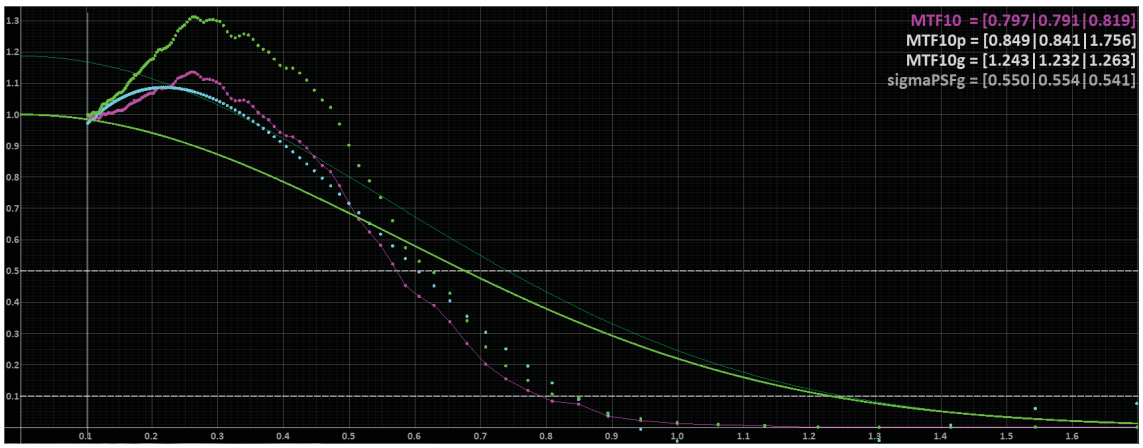


Figure 4.8. Measurements and Approximation for an image under influence of sharpening artefacts. CTF measurements (green dots), MTF measurements (magenta dots), Gaussian fit (green bold), Polynomial fit (cyan dots), Piecewise linear interpolation (magenta line segments)

predicted values and measurement values has come below a minimum error threshold (e.g. $e_t = 0.01$). Then, the resolution limit is found by solving the Gaussian model function for k_{MTF10}.

4.3.2 Polynomial Fit

Another way to approximate MTF is to fit a polynomial function into obtained measurements (see Figure 4.8). There, the polynomial coefficients according to a function with predefined degree (e.g. $n_p = 5$) are found as part of a lower–upper (LU) matrix decomposition and factorization. Complexity of LU-decompostion usually is smaller than QR-decomposition and thus preferable [77].

Polynomials with higher order can not be directly solved for k_{MTF10}. Therefore, resolution limit frequency k_{MTF10} is found as an iteration through a vector filled with values for k_i and a small step size (e.g. 0.1Hz) and corresponding values for polynomial $\tilde{H}(k_i)$. The while-iteration stops, when $\tilde{H}(k_i) < 0.10$ followed by an interpolation between $x_1 = k_{i-1}$, $y_1 = \tilde{H}(k_{i-1})$, $x_2 = k_i$ and $y_2 = \tilde{H}(k_i)$.

4.3.3 Piecewise Linear Interpolation

A simple and straightforward approach to approximate MTF is piecewise linear interpolation between every MTF value $\tilde{H}(k_i)$ and its predecessor $\tilde{H}(k_{i-1})$. The shape of this MTF approximation precisely follows previously determined CTF (see Figure 4.8) since both functions are directly linked by a normalization series (see equation 3.7). Similar to polynomial fit, the resolution limit frequency k_{MTF10} is found as result of a while-loop until $\tilde{H}(k_i) < 0.10$ followed by linear interpolation between $x_1 = k(i-1)$, $y_1 = \tilde{H}(i-1)$, $x_2 = k(i)$ and $y_2 = \tilde{H}(i)$. Thus, the obtained

function is continuous but the derivative is not, despite the fact that any order derivatives of MTF are rarely needed. If continuous derivatives become necessary the piecewise linear method can be substituted for an bi-cubic approach.

4.3.4 Validation

On the one hand, both fitting methods, Gaussian and polynomial, deliver functions with known derivatives in the whole range of definition but it should be mentioned again that any order of MTF derivative is rarely needed. On the other hand, both approaches use all measurements as input and if these are falsified, e.g. due to inhomogeneous target illumination or sharpening artefacts, approximated values may vary tremendously (see Figure 4.8).

Given the assumption that values are obtained correctly with the presented approach in section 3.3 the main advantage of piecewise linear interpolation is that no general mistakes are made as it occurs with Gaussian and polynomial fitting procedures. MTF measurements of one hundred images and six different cameras in relation to applied approximation approach and differences to piecewise-linear method can be found in Table 5. Mean difference between polynomial fit and linear approximation is 4.8% (red 5.1%, green 2.9%, blue 6.5%). The difference between Gaussian fit and linear approximation is much larger. There, the mean deviation is 15.6% (red 15.5%, green 11.9%, blue 19.3%).

One way to diminish the influence of falsified measurements for both fitting approaches is to limit the data-range around an expected resolution limit k_{exp}. A further way for Gaussian fit is to expand the function with Hermitian polynomials [8].

Under the premise that the resolution limit is obtained as k_{MTF10} and the piecewise linear method is unresponsive to falsified measurements at lower spatial frequencies, this shall be the default method for MTF approximation.

4.4 Siemens-Star Center Position

The developed algorithm in section 3.3 commences by scanning the image in circles having one center in common and presuming correct position. Thus, determination of Siemens-star center position is an important step [78, 79].

In this thesis, three different techniques for center extraction are considered and described. It can be shown that all three approaches deliver very similar center position values and can me used exchangeable or combined while supporting each other's measurements.

4.4.1 External Markers

The target layout, introduced in section 3.3.2, includes 4 rotor-shaped markers which often can be found as ground control points in geodetic and photogrammetric survey. The marker consists of two opposing white segments and two opposing black segments

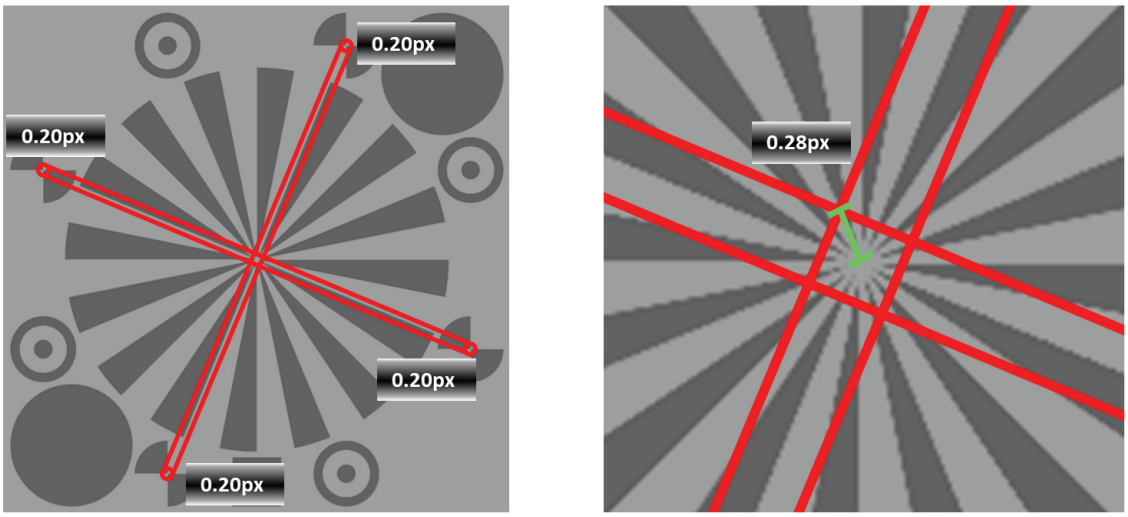


Figure 4.9. Siemens-star center determination with external markers

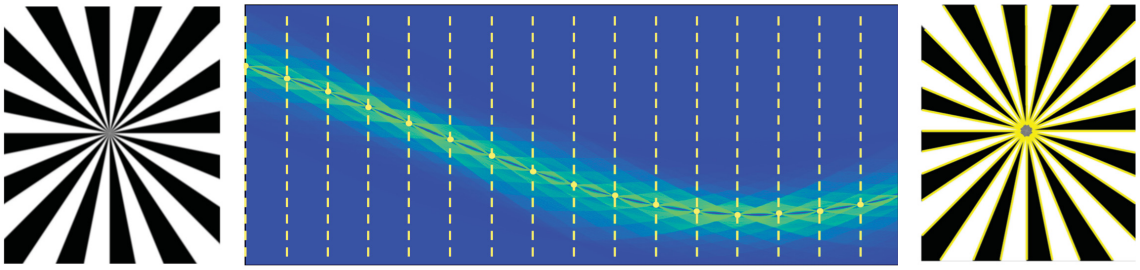


Figure 4.10. Siemens-Star line segmentation by hough transformation and peak detection

perpendicular to the aforementioned. Each external marker center can be picked manually with an accuracy better than 0.2 pixel [67].

The star center itself is calculated as intersection of the two lines spanned by opposing markers (see Figure 4.9). Applying error propagation for both lines and related single marker accuracies (< 0.2 pixel) the center can be calculated with an accuracy better than 0.28 pixel.

4.4.2 Line Segment Detector

The target layout with it's square wave pattern (as opposed to a sinusoidal wave) also qualifies to use the edges of each adjacent black and white segment to determine the star center position. Two different line detection approaches will be introduced and discussed hereafter.

The standard Hough transform is often used to detect edges in images (see Figure 4.10). There, straight lines of the form $y = mx + n$ are represented as the Hesse

normal form instead of a single point, e.g. (m, n), to prevent problems with vertical lines [80]:

$$r = x \cos\theta + y \sin\theta \tag{4.14}$$

Main advantage of standard Hough transform is the tolerance towards gaps between line segments and noisy input images. Disadvantage is that detected lines are infinite, described by their (r,θ) values, rather than finite lines with defined start and end points.

Another method for line segment extraction was introduced by Gioi et al. [81] based on Burns et al. [82]. In contrast to standard Hough transform this approach delivers finite lines with defined start and end points.

The algorithm "is aimed at detecting locally straight contours on images" and the concept of contours is based on "zones of the image where the gray level is changing fast enough from dark to light or the opposite" [81]. The line, perpendicular to the direction of gray level change, is called "level line" and it points in direction of the later extracted line segment. "The algorithm starts by computing the level-line angle at each pixel to produce a level-line field, i.e., a unit vector field such that all vectors are tangent to the level line going through their base point. Then, this field is segmented into connected regions of pixels that share the same level-line angle up to a certain tolerance τ. These connected regions are called line support regions.". A subsequent validation stage decides if a line support region is considered a line segment or not.

It should be noted that the algorithm "is designed to work on any digital image without parameter tuning" and "controls its own number of false detections". The Siemens-star center position is determined by applying line segmentation to the test pattern and calculating the mean of all possible intersection combinations for the extracted line segments. However, a single calculation of all intersections and correspondent mean value slightly differs depending on chosen region of interest and thus depended on initial assumption of center position (e.g. operator input). The slight differences assumingly occur due to the algorithm's self controlled parameters. To find center position more precisely, several measurements in close vicinity (e.g. ± 3 pixel for both image axis x,y) of the initial center are obtained and averaged. It can be shown, that the statistical distribution is clustered around the real center position (see section 4.4.4).

4.4.3 Phase Shift Approach

Impact of falsely determined Siemens-star center on MTF measurement has been investigated by Birch and Griffin [78] (see Figure 4.11). There, an equation reflecting a constant offset has been introduced for sinusoidal Siemens-stars. After fitting the scanned intensity values to a sine function, aforementioned equation then can be solved for offset coordinates in both directions (x,y). However, this approach is not well suited for square wave patterns.

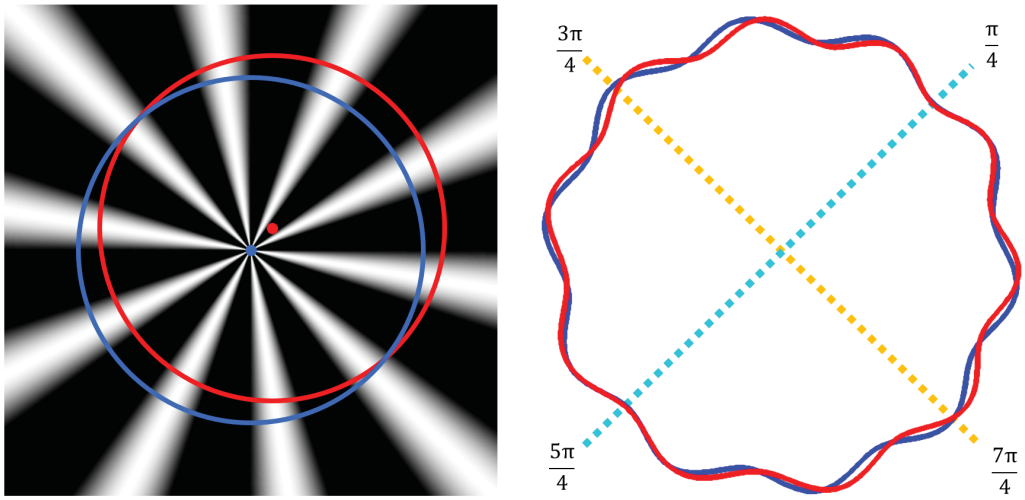


Figure 4.11. (left) Correct centered (blue) and falsely centered (red) circles, (right) plot of related scanned intensity (image inspired by Birch and Griffin [78])

The presented approach in section 3.3 is designed to work for both square and sinusoidal Siemens-stars. In the latter case, normalization according to Coltman [66] must be omitted.

The method used in this present work to exploit phase shift for center detection is structured as follows. Test pattern frequency, as the number of black-white segments, is determined as frequency (and phase) belonging to the sine wave having the largest magnitude contributing to the square signal after Fourier transform of the outermost circle (low spatial frequency, see Figure 3.4 and section 3.3). Inverse Fourier transform (only for the dominant frequency) then delivers a sinusoidal reconstruction of the square wave. As a side note, phase of this reconstructed sine wave most likely is different to phase at the correct center position because shape and phase of the false (de-centered) square signal scan is not invariant to spatial shift of the image coordinates. Aim is to determine phase shift between the reconstructed sine wave and the scanned (extracted) signal. This can be done by locating intensity minima and maxima for both signals and compare their related shift. These minima and maxima are found by applying a peak detection technique to either signal and compare their related position (angle, see Figure 4.12). Extrema are obtained by validating if the sign of calculated slope changes. This procedure is being applied at every position of the circle:

$$\text{Maxima x(i): } [x(i) - x(i-1)] > 0 \text{ and } [x(i+1) - x(i) < 0] \tag{4.15}$$

$$\text{Minima x(i): } [x(i) - x(i-1)] < 0 \text{ and } [x(i+1) - x(i) > 0] \tag{4.16}$$

For robust implementation, the span in x-direction can be extended (e.g. to a ± 3 neighbourhood) and further a minimum amplitude difference of possible extrema

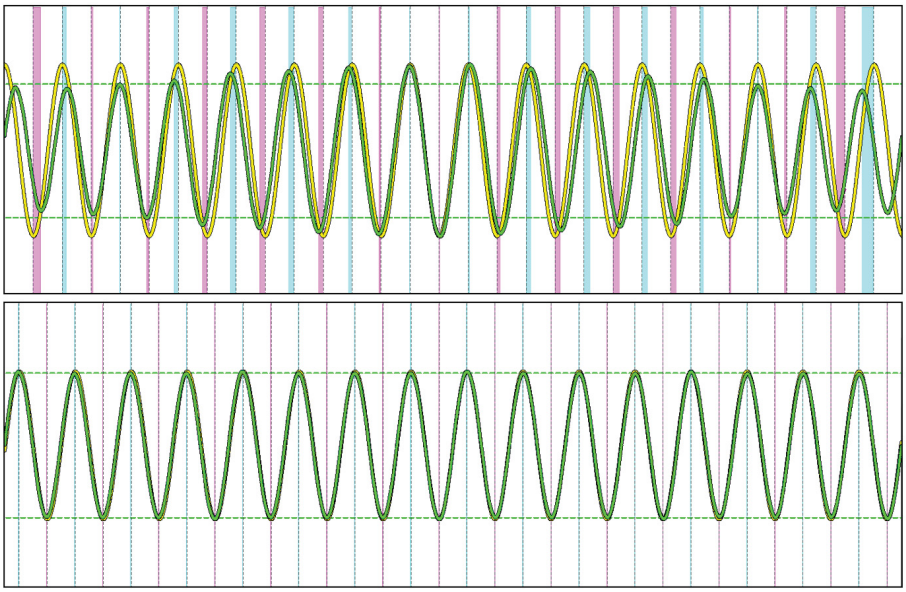


Figure 4.12. (top) Phase shift between de-centered circle and reconstructed sine signal, (bottom) phase shift between centered circle and reconstructed sine signal

compared to overall signal mean can be introduced.

However, peak detection for square signals is not unambiguous because slope at minima / maxima position and their larger surrounding is equal to zero which violates the requirements of equation 4.15 and 4.16. To overcome this issue, following observation is very useful. Scanned signals of square Siemens-stars become more and more sinusoidal shaped when modulation contrast falls below 1 (see Figure 4.13) and peak detection can be applied. Therefore, an additional constraint as upper bound is formulated so that peak detection and phase shift only is calculated if modulation contrast is below a certain threshold (e.g. $b_u = 0.9$, $\tilde{H}(k) < b_u$). When spatial frequency k approaches resolution limit the extracted signal will tend to a flat line and peak detection again is not unambiguous. As a consequence a lower bound is set where peak detection and phase shift can be calculated (e.g. $b_l = 0.2$, $\tilde{H}(k) > b_l$). Final result for phase shift S at a particular Siemens-star position (x,y) then can be formulated as a double sum of all phase shifts (for all extrema) and every circle j between upper and lower bound b_u, b_l, where a single phase shift is the distance between reconstructed peak p_r and scanned peak p_s:

$$S(x,y) = \sum_{j} \sum_{Extrema} dist(p_s, p_r) \tag{4.17}$$

Having an initially determined center (e.g. given by an operator) the previously described method is performed at all positions in a predefined neighbourhood (e.g. ± 3 pixel for x,y) with sub-pixel step size. The position (x,y) corresponding to the

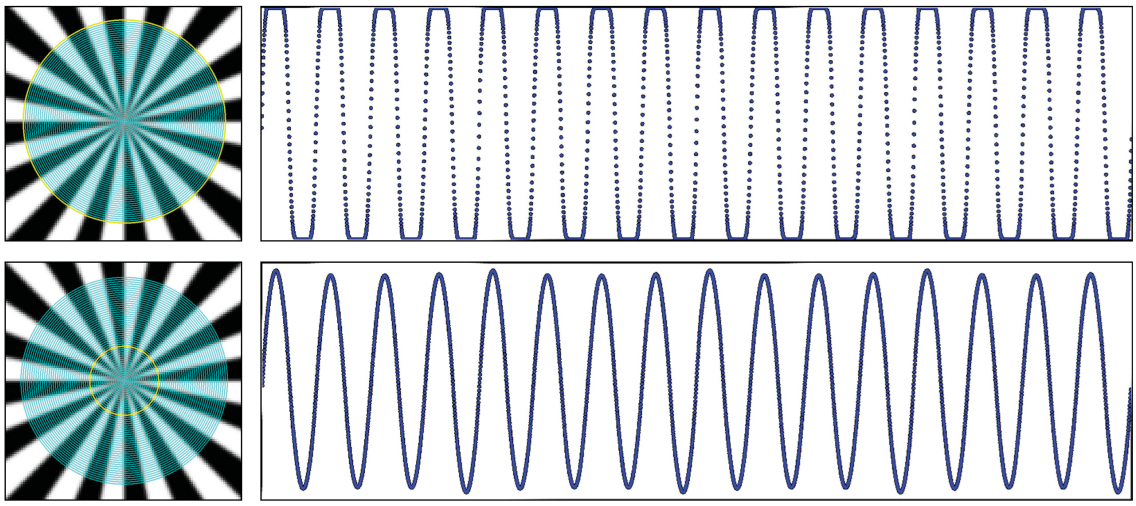


Figure 4.13. (top) Scanned circle (yellow) at lower spatial frequency and square shaped signal scan, contrast modulation equal to 1, (bottom) scanned circle (yellow) at higher spatial frequency and sinusoidal shaped signal scan, contrast modulation smaller than 0.9

minimum phase shift sum S then is the correct Siemens-star center position. An exemplary phase shift distribution is depicted in Figure 4.14 (lower left).

As a final side note, using a gradient descent approach to find the minimum in aforementioned distribution is much faster than scanning the entire neighbourhood but at the risk of ending up in a local minima. With the proviso to define a standardized method real-time measurements should not be priority but correct and outlier free results.

4.4.4 Validation

Three different approaches are available to determine the correct Siemens-star center position, using a) line segments b) signal phase shift and c) external markers if included in the test target. To validate all three methods, twenty images have been selected showing the proposed design of the star to which every technique then has been applied. Table 6 shows the maximum distance between each approach having an average of 0.21 pixel. A further way to illustrate the relation is to calculate the mean of all three techniques (per image) and then plot the related distance for every method (see Figure 4.14, top). The maximum distance to the averaged center coordinates is 0.21 pixel and the mean distance 0.08 pixel.

To clarify the impact of these center-location differences, spatial frequency values k_{MTF10} have been calculated at every position and subsequently differences between minimum and maximum value have been determined (see section 8.2.5, Table 7). Thereby, mean deviation of 2.2 % can be considered very low.

An alternative comparison is shown in Figure 4.14 (middle) as deviation in percent to the related average of k_{MTF10}. Again, mean deviation values for phase shift (+0.60%), external markers (−0.21%) and line segments (−0.44%) is very low.

Another confidence interval can be given as the distribution of the MTF itself in close vicinity of the star center, e.g. calculated as mean of all available center refinement approaches (see Figure 4.14, lower right). Birch and Griffin [78] already observed: "Small pixel shifts in the center location are shown to induce significant degradation in the measured SFR compared to the nominal, correctly centered measurements.", where they explicitly use the term "degradation" instead of variation leading to the conclusion that the maximum MTF10-value occurs at the correct star center position. Therefore, it is useful to add that value and corresponding center position (x, y) to the measurement report.

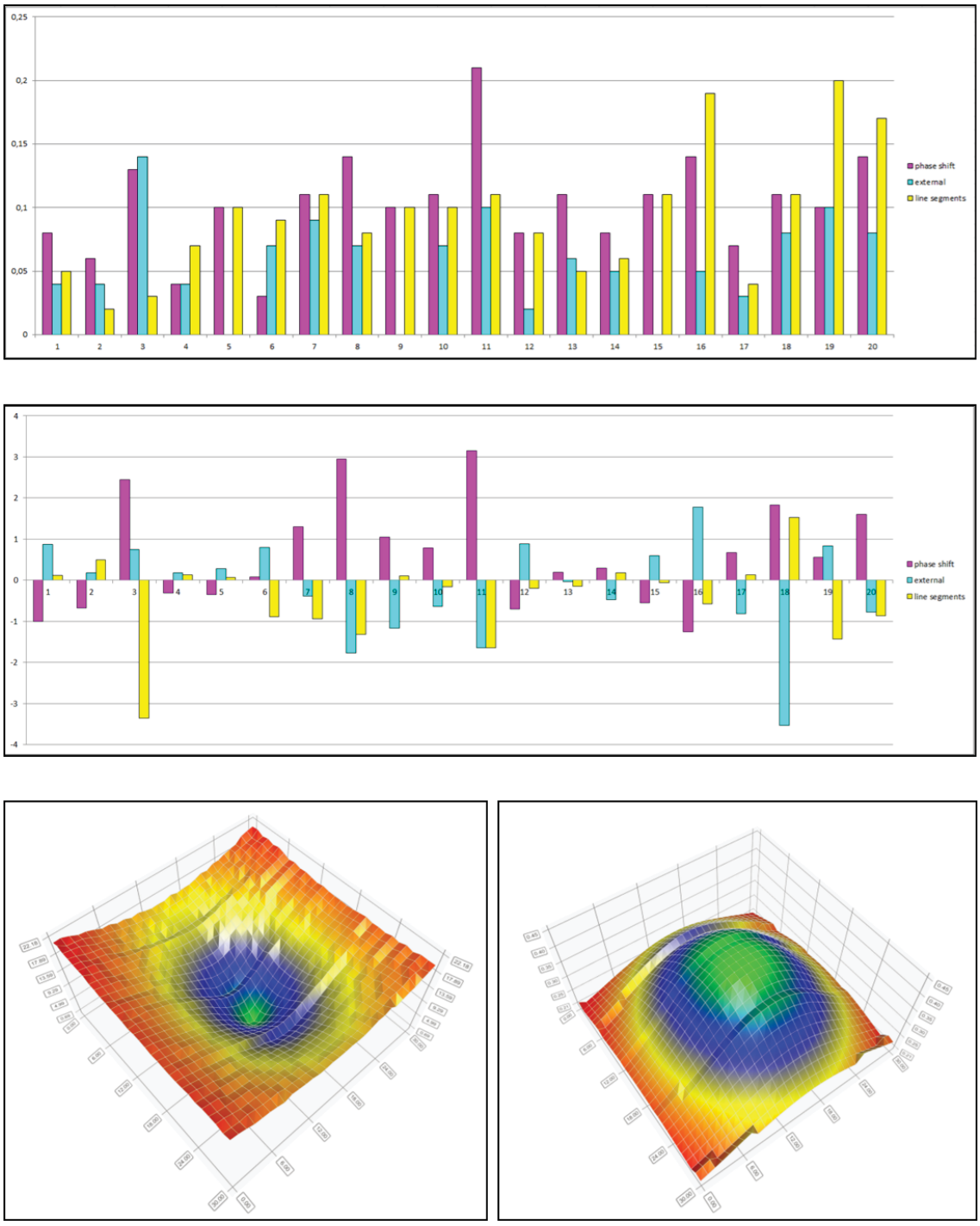


Figure 4.14. (top) , (middle) , (lower left), (lower right)

4.5 Normalization of Contrast Magnitude

An essential step of the approach using Siemens-stars to determine spatial resolution is the normalization to contrast level C_0 at spatial frequency equal to 0 (infinite radius, see equation 3.5 in section 3.3.1). Usually, this is done by calculating C_0 as the amplitude between intensities $I_{max}(k)$ and $I_{min}(k)$ for the outermost circle where $k = k_0$:

$$C_0 = \frac{I_{max}(k_0) - I_{min}(k_0)}{I_{max}(k_0) + I_{min}(k_0)} \tag{4.18}$$

According to equation 4.18, C_0 only is equal to 1 and thus neglectable if $I_{min}(k_0)$ is equal to 0 which in turn is rarely the case (depending on material of the star and light conditions during image acquisition). If $I_{min}(k_0)$ is not equal to 0 and C_0 is not considered spatial resolution will be determined too optimistically (too good). For that reason normalization is an essential step.

4.5.1 Specified Arrangement

Determination of spatial resolution using the proposed 32 segment Siemens-star layout (see section 3.3, Figure 3.5) when printed out e.g. with 135 x 135 cm^2 (related star diameter 102cm) will be limited to 3.5cm GSD without using external normalization patterns. If this GSD can not be kept (is larger) it can in turn not be guaranteed that the outermost ring contains pure black and white values and the use of external normalization patterns becomes mandatory.

It is then theoretically sufficient to have one scanned circle slightly above MTF at 10% modulation contrast level and one below to determine spatial resolution limit. That would extend the theoretical measurable GSD up to 6.0cm or even 7.0cm. But this in turn lowers the capability to determine correct star center position (see section 4.4). External markers are to small an therefore can not be used any more, line segment's length is only 5.0 - 6.0 pixel an thus prone to failure and similarly phase shift approach also has only 5 - 6 circles to determine the correct center. Taking all considerations into account it is expedient to use a layout where one star segment of the outermost ring is approximately three times larger than the intended GSD (3.5cm for this example, test pattern size 135 x 135 cm^2).

As a final note, aforementioned considerations and calculations are based on GSD instead of GRD (the real pixel footprint contributing to the signal). If GRD is significantly worse compared to GSD, e.g. $k_{MTF10} = 0.700$ [line/pixel] and thus GRD = 5.0cm (equation 3.2 with GSD = 3.5cm) either the layout has to be enlarged or flight height needs to be lowered. But having an estimate of GRD is not trivial because this is exactly the value that is going to be determined as spatial resolution. If feasible, a priori laboratory measurements (see section 5.7) could deliver a first estimation.

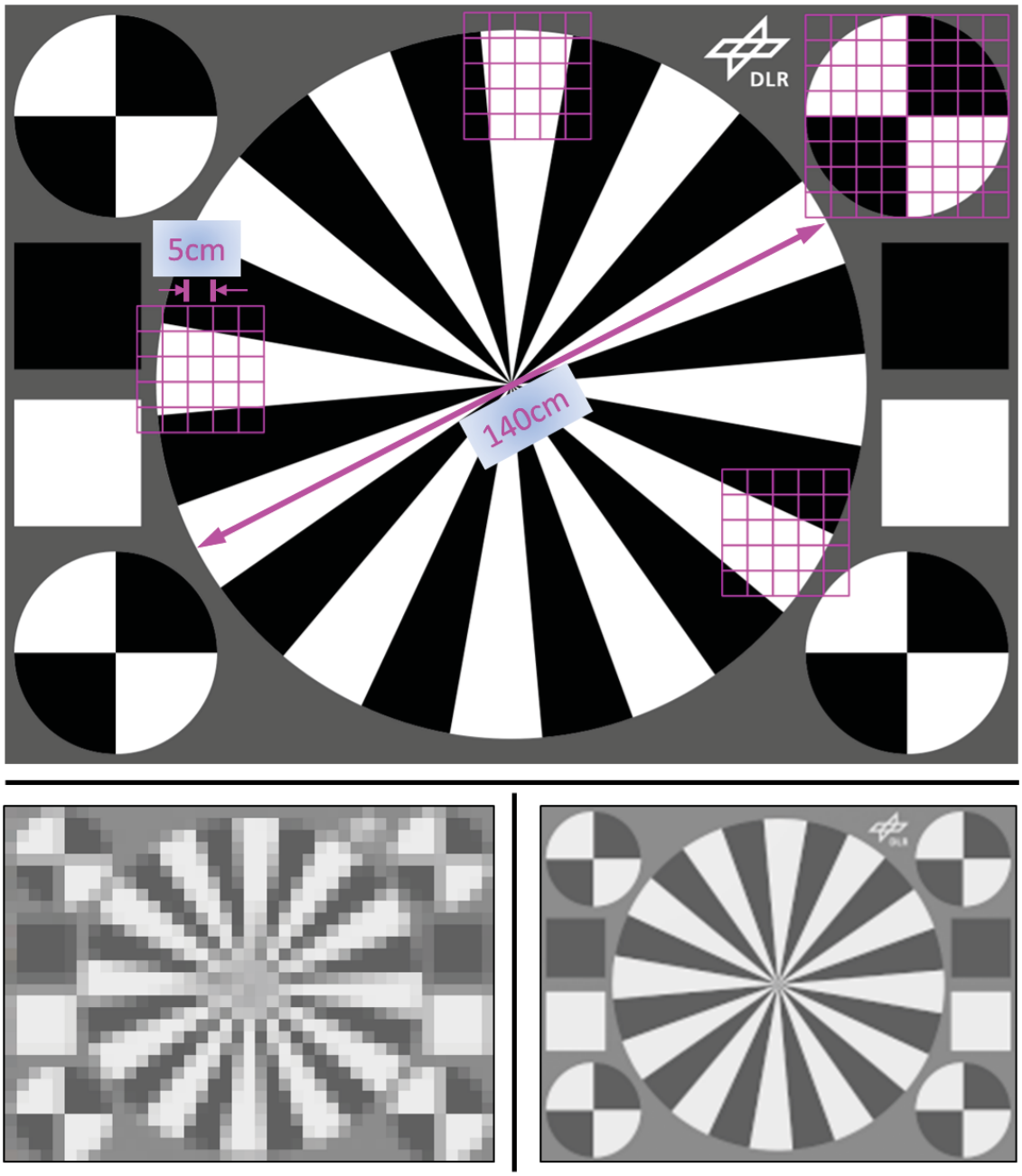


Figure 4.15. (top) Proposed alternative Siemens-star layout with 12 black-white segments, (bottom left) simulated GSD of 5.0cm zoom factor 1500% measured spatial frequency $k_{MTF10} = 1.327$ line/pixel, (bottom right) simulated GSD of 1.0cm zoom factor 300% measured spatial frequency $k_{MTF10} = 1.287$ line/pixel

4.5.2 Validation

As many aerial camera systems, especially non-UAV versions, are mounted to (mostly) forward gliding aircrafts having a minimum flight speed, minimum flight height to avoid motion smear (for more detail see section 4.7) and minimum camera exposure time, genuine 3.5cm GSD might not be feasible to deliver. Therefore, the resolution test target should be able to provide the basis for correct measurements of at least 5.0cm or better between 8.0cm and 10.0cm GSD [83, 84] while being mobile (transportable) for flexibility.

An alternative design can be seen in Figure (4.15, top). One edge is stretched to obtain enough space for external markers. If the presented design is printed out in 150 x 200 cm^2 (star diameter 140cm, marker diameter 40cm) and then captured at an intended GSD of 5cm the external markers are imaged in a 8 x 8 neighbourhood which is (empirically) the bare minimum to obtain precise marker center coordinates ans subsequently star center coordinate.

Furthermore, the alternative test pattern is constructed of 24 segments instead of 32 segments which is a compromise between having sufficient directions for motion detection (or 2D-MTF-measurement in general) and sufficient space between the segments so that the blur circle's diameter is smaller (compared to more segments) and thus ensures that line segments consist of more pixels and the phase-shift approach has more circles to robustly determine correct star center position if external markers can not be used anymore. The normalization patterns are adopted to intended GSD and placed redundantly on the target to enable minor noise estimation.

To validate the layout and verify if the intended GSD leads to measurable results the design has been rendered using a synthetic scene and a simulator for validating photogrammetric algorithms [85]. The results are presented in Figure (4.15, bottom left and bottom right). It can be seen that intended GSD of 5.0cm is the maximum where external markers can be used. Measured spatial frequencies $k_{MTF10} = 1.327$ for 5.0cm GSD and $k_{MTF10} = 1.287$ for 1.0cm GSD only differ by approximately 3% and can therefore be considered almost similar determined.

Concluding with several considerations regarding test target size and included normalization patterns, the compromise between a large range of (standardized) measurable resolutions and flexible target deployment has to be made depending on individual circumstances. While for some end-users a target size of 135 x 135 cm^2 is too large others might be in a position to deploy a 200 x 200 cm^2 version which then would easily enable measurements of GSD up to 10.0cm. Splitting the target into several sub-components (e.g. 4 pieces à 100 x 100 cm^2) only is an option if the manufacturer guarantees seamless print and fixture. Another alternative, printing the design on fabric or tarpaulin, entails new prerequisites. Then it has to be ensured that the target, when deployed for measurement, is completely flat and free of folds.

When aiming for standardized measurements for large scale aerial cameras (in contrast to UAV-based cameras) it is probably the best solution if sufficiently large and

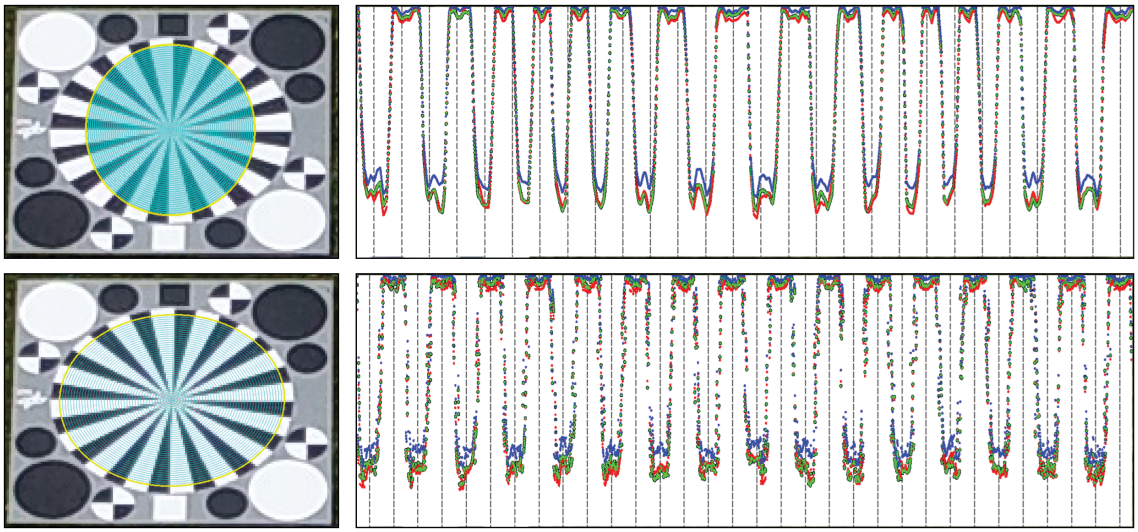


Figure 4.16. (top) Imaged Siemens-star, scan coordinate system and signal scanline without homographic transformation (bottom) imaged Siemens-star, scan coordinate system and signal scanline with applied homographic transformation

appropriate targets are provided (stationary) by e.g. major research institutions, government agencies or even large private companies as calibration service for their customers.

4.6 Influence of Test Target Inclination

In aerial photogrammetry there are several circumstances where the Siemens-star test target is imaged inclined. For example close to image borders if camera-lens combinations with large aperture angles are used or if oblique camera setups are utilized in general. Then the test pattern is imaged including projective distortion (see Figure 4.16). This kind of inclination affects spatial resolution measurement in terms of finding correct extrema and subsequently obtained contrast and modulation magnitude. If the circles are scanned without projective transformation then occasionally minima occur at maxima positions and vice versa (see Figure 4.16, top) which in turn leads to mostly too pessimistically (worse) determined spatial resolution values.

In computer vision the transformation between two projective planes can be achieved by applying a 2D homography $\mathcal{H}$ [86, 87] which transforms a pair of coordinates (x,y) from one plane to a corresponding pair of coordinates (x',y') to the other plane (see following section 4.6.1). As a side note, formula symbol for the 2D homography is chosen to be $\mathcal{H}$ in this present work instead of H, because this is the reserved symbol for the point spread function (equation 2.3).

4.6.1 Coordinate System Transformation

The mathematical conception to determine a 3×3 - homography matrix $\mathcal{H}$ having eight degrees of freedom for $n \geq 4$ homologous image points $x_i \leftrightarrow x_i'$ according to Rodehorst [87] is defined as follows:

$$x_i' = \mathcal{H} x_i \tag{4.19}$$

To obtain a unique solution, points included in the calculation neither must be distributed co-linear nor identical. After having the image points transformed to homogeneous coordinates a normalization is advised to achieve numerical stable results. For that purpose, all points $\mathbf{x}_i = (x_i, y_i, w_i)^T$ are shifted in a way that the common centroid is identical to the point of origin $(0, 0)^T$. Subsequent point scaling is applied to adjust a mean distance to origin of coordinates equal to $\sqrt{2}$. Following transformation matrix includes both translation and scaling:

$$T = \begin{bmatrix} \lambda_x & 0 & t_x \\ 0 & \lambda_y & t_y \\ 0 & 0 & 1 \end{bmatrix} \tag{4.20}$$

Then, the normalized coordinates result from corresponding multiplication $\hat{x}_i = T x_i$. The normalization must also be calculated for the second (projectively rectified) image to transform image points x_i' with T' to $\hat{x}_i'$. For all (now normalized) image coordinate pairs $\hat{x}_i \leftrightarrow \hat{x}_i{}'$ the two lines for the partial design matrix $\mathcal{A}_i$ must be set up:

$$\mathcal{A}_i = \begin{bmatrix} \hat{w}_i' \hat{x}_i^T & \vec{0}^T & -\hat{x}_i' \hat{x}_i^T \\ \vec{0}^T & \hat{w}_i' \hat{x}_i^T & -\hat{y}_i' \hat{x}_i^T \end{bmatrix} \tag{4.21}$$

Then, all observations are summarized in a common $2n \times 9$ matrix. Having $n \geq 4$ homologous data samples, the linear homogeneous system of equations

$$\mathcal{A} h = 0 \tag{4.22}$$

with $h = [h_1, ..., h_9]^T$ can be solved by a singular value decomposition (SVD). The solution vector h corresponds to the eigenvector belonging to the smallest eigenvalue of $\mathcal{A}$. When the components of the solution vector h are brought back into matrix form, the homography matrix is obtained for the normalized coordinates:

$$\hat{\mathcal{H}} = \begin{bmatrix} h_1 & h_2 & h_3 \\ h_4 & h_5 & h_6 \\ h_7 & h_8 & h_9 \end{bmatrix} \tag{4.23}$$

In order to obtain the homography matrix for the original coordinates, the normalization must be reversed. This is done by:

$$\mathcal{H} = T'^{-1} \hat{\mathcal{H}} T \tag{4.24}$$

4.6.2 Validation

The required $n \geq 4$ homologous image points $x_i \leftrightarrow x_i'$ will be provided by the proposed test pattern itself. The (operator-driven) image coordinates of the external marker correspond to x_i' and precise knowledge of the test target (metric object space coordinates) provide the counterpart set x_i. The obtained solution for the 3×3 - homography matrix $\mathcal{H}$ is defined only up to scale (nine matrix elements but eight degrees of freedom). However, in aerial photogrammetry the necessary scale factor is equal to ground sample distance and thus metric, compatible to object space and can be calculated from pixel-size, focal length and flight altitude (equation 3.2).

The homographic transformation can be applied in both directions. One way is to rectify the original (distorted) image and scan the undistorted image as described in section (3.3). Disadvantage of this procedure is that it includes an image re-sampling step which affects spatial resolution and its measurement. Therefore, the other direction is advised. With known homography matrix $\mathcal{H}$ the entire measurement coordinate system S_{xy} (equation 4.2) is transformed according to test pattern inclination (see Figure 4.16, bottom left).

Accuracy of the proposed method to correct test target inclination relies on two circumstances. First, precise point measurement of the external markers (approx. 0.2 pixel [67]) and second correct scale factor determination and thus dependent on precise knowledge of exterior and interior orientation of the camera system.

As both parameter sets almost never will be free of error, calculation and utilization of the homography matrix will unavoidably induce errors to the scanned circles. While obtained signals now have the correct peak positions the overall shape becomes scattered (see Figure 4.16, bottom right) with rising manifestation for larger spatial frequencies which in turn leads to sometimes over- or underestimated extrema. Due to this fragile behavior two suggestions shall be provided. If oblique camera systems are investigated / calibrated the used test targets should ideally be placed inclined on the ground to correct the oblique camera angle to a quasi-nadir setting. If this is not possible or cameras with large aperture angles are under investigation the overall measurement confidence has to be decreased. Empirical observations, including all standardization aspects from previous sections, suggest an average measurement accuracy of approximately 3% - 5%. If a homographic transformation is applied this confidence level has to be changed (e.g. 5% - 8%) and obligatory mentioned in the measurement protocol.

4.7 Motion blur

Sensor motion of aerial imaging systems is either caused by sensor-rotation with three DOF (roll, pitch and yaw of the aircraft reference frame), sensor-translation with three DOF (X Y, Z, e.g. world coordinates) or due to both motions at the same time. An optical sensor under motion during image acquisition not only collects photons of the static pixel-footprint projected onto the observed surface but of the

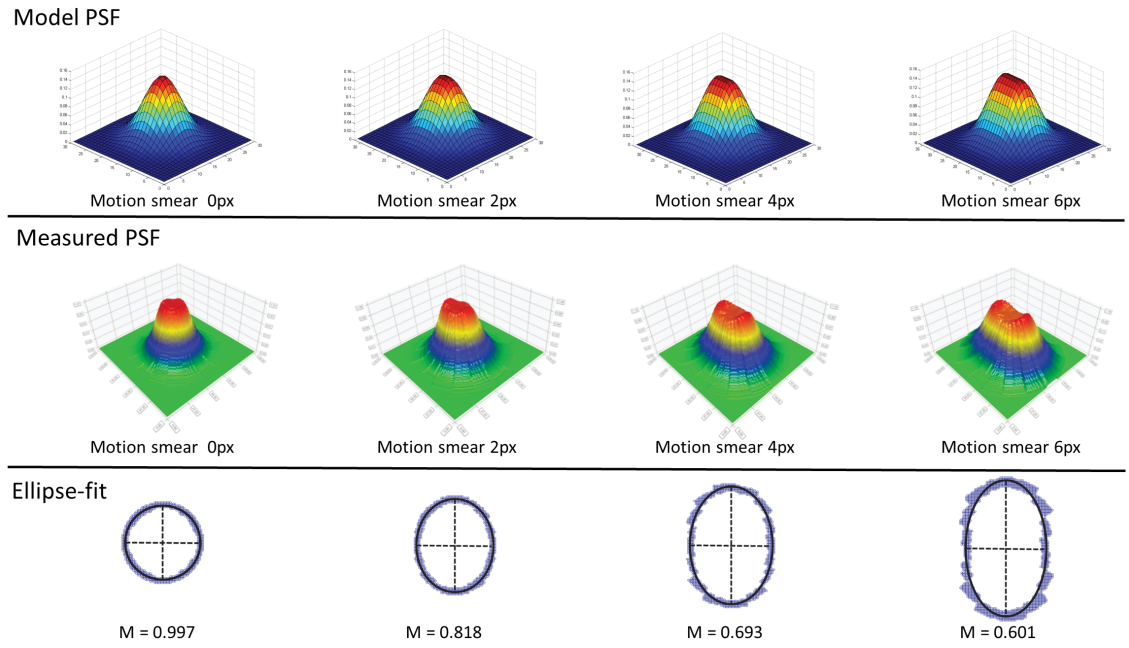


Figure 4.17. (top) model convolution-kernel with rising influence of motion from 0 to 6 pixel, (middle) measured PSF of images convoluted with model-kernels, (bottom) corresponding ellipse fit for 2D-MTF10 and relation M of semi-minor ellipse axis e_s and semi-major ellipse axis e_l

extended footprint along the projected line of movement. Then, the input signal $U(x', y')$ (equation 2.3) can be described as an integral along the projected line of movement in object space:

$$U(x', y') = \int U_\sigma(x', y', m)\ dm \tag{4.25}$$

Where $U_\sigma(x', y', m)$ is the input signal at every projected position depending on motion m (6 DOF) during the exposure time window.

During that window motion induces smear which directly affects the overall PSF $V(x, y)$ (equation 2.3). With increasing motion (e.g. 1D-translation in flight direction of an aircraft at high velocity and/or long exposure time frame) the PSF will be smeared along the projected motion.

4.7.1 Characterization

The aforementioned smear along the projected motion can be characterized and investigated empirically using a simulation setup. A similar simulation is later used again to compare different measurement techniques (Slanted-edge and Siemens-star approach, see section 5.5).

The sequence is to apply predefined modulation (MTF) or spread parameters (PSF) to an ideal representation of resolving patterns (see Figure 5.5). This can be done by forming a convolution of mathematical-ideal image-intensity values of an image (I), a Gaussian-shape model PSF (H_m) and a mathematical-ideal sensor PSF (H_s). Simulated PSF (H_{sim}) then can be formulated as follows:

$$H_{sim}(r) = I(r) * H_m(r) * H_s(r) \tag{4.26}$$

Convolution of image-intensity values $I(r)$ and an increasingly smeared kernel (initial) Gaussian-shape (see Figure 4.17, top) then delivers a more and more stretched version of corresponding measured PSF $H(r)$ (see Figure 4.17, middle).

Measured PSF of a motionless kernel is circle-shaped and becomes increasingly elliptic-shaped when smear gains influence (see Figure 4.17, bottom). Therefore, the relation between length of semi-minor ellipse axis e_s and semi-major ellipse axis e_l can be used as measure M of induced motion:

$$M = \frac{e_s}{e_l} \tag{4.27}$$

Then, circle-shaped 2D-PSFs and corresponding ellipse fit will deliver values for M close to 1.00 and decreasing values for increasing elliptic-shape of measured 2D-PSF (see Figure 4.17, bottom). This measure will be used to evaluate dominant motion smear of aerial images (see following section 4.7.2).

4.7.2 Validation

A general comparison of aerial images with and without dominant motion smear is given in Figure 4.18. The image in the upper row suffers from sharpness-loss due to motion smear. This becomes clear when looking at the grade of 2D-PSF-deformation (Figure 4.18, top right). Quotient M according to equation 4.27 in this case is 0.680. The image in the lower row does not suffer from dominant motion smear and the related Quotient M in that case is 0.967.

A practical application for motion blur detection is to investigate if planned and executed flight campaigns deliver sharp aerial images without dominant smear. Generally the flight plan contains: intended flight altitude (and thus intended GSD), in track and across track overlap (according to aircraft cruising speed) and intended exposure time which gives a first estimate of expected motion smear.

Usually, fight campaigns for 3D mapping or 3D reconstruction are planned with 80% in track overlap of and 60% across track overlap [88]. Then, the Siemens-star test target is imaged at 15 different locations in image space when maintaining predefined flight plan. Using the presented software-tool for spatial resolution the 2D-PSF and corresponding quotient M (relation of semi-minor and semi-major ellipse axis) can be calculated for every position.

Exemplary results for a benchmark campaign using the high precision UAV test field

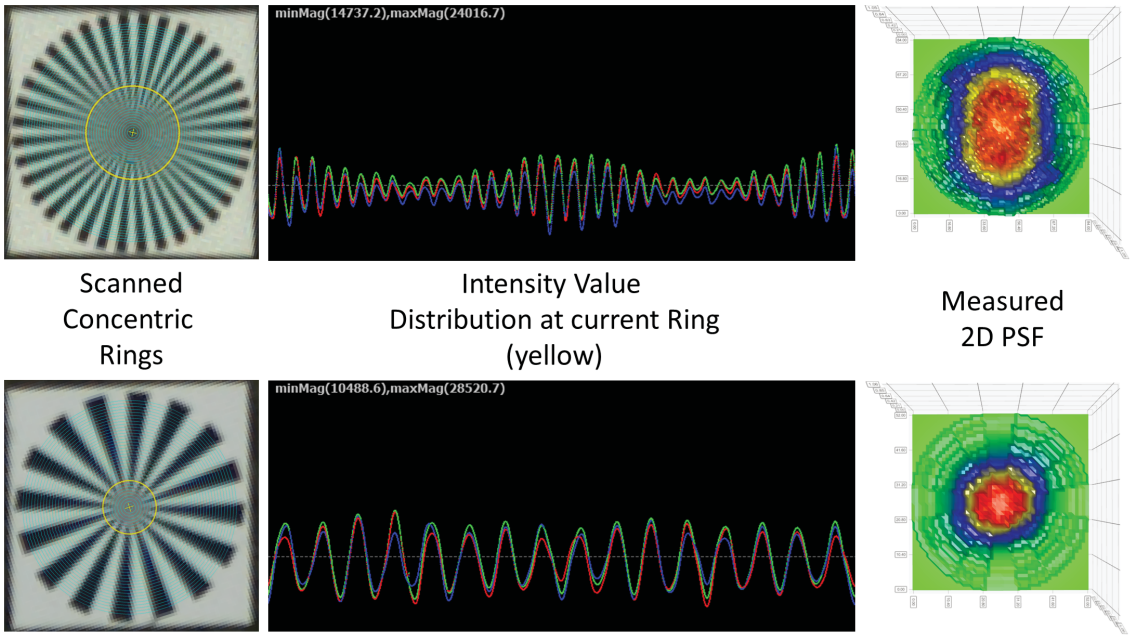


Figure 4.18. (top) UAV image with dominant motion smear, (bottom) UAV image without dominant motion smear

Zollern Colliery [89] are given in Table 8. Mean value of all 15 images is $M = 0.912$ and reveals that the length difference between semi-minor and semi-major axis is very small. Furthermore, it strongly indicates that 2D-PSF is rather circle-shaped and leads to the conclusion that there is no dominant motion smear present for this particular flight campaign.

As mentioned earlier in this section, motion has six degrees of freedom. The approach of using the quotient of semi-minor and semi-major ellipse axis will not be applicable if the dominant forward motion is coupled with e.g. a rotation around the roll axis of the aircraft reference frame. Then, 2D-PSF and corresponding quotient could indicate a circular shape leading to the conclusion that no motion blur is present but concealing at the same time that the roll-rotation deteriorated spatial resolution perpendicular to the flight direction. It is therefore beneficial to have an estimate for spatial resolution under static (laboratory) conditions and under dynamic (operating) conditions (see section 5.7 and 5.8 for in depth discussion).

Having an initial estimate for static and dynamic conditions (and their relation) supports the decision making process if dominant motion affected the images or if a coupled motion is inherent in the images, assuming all other standardization aspects explained in this chapter are fulfilled. A measurement of circle shaped 2D-PSF but significantly worse spatial resolution compared to the (calibrated) dynamic estimate then would indicate a complex motion during the exposure time frame.

4.8 Conclusions

In this chapter seven different aspects have been explained and validated in preparation of DIN 18740-8 and a standardized measurement of spatial resolution for optical remote sensing systems:

- Demosaicing methods (section 4.1)
- Signal scan interpolation methods (section 4.2)
- MTF approximation / interpolation (section 4.3)
- Siemens-star center position determination (section 4.4)
- Normalization of contrast magnitude (section 4.5)
- Influence of test target inclination (section 4.6)
- Influence of motion blur / smear (section 4.7)

This section shall provide a compact summary of all findings, conclusions and suggestions regarding objectively determined spatial resolution.

Demosaicing methods: The utilized demosaicing method needs to be mentioned (logged) in the measurement protocol to enable classification of obtained results and due to significant performance differences between the presented approaches. In aerial remote sensing scenarios it is further advised to only use AHD approach (section 4.1.2) if a robust outlier detection during bundle block adjustment can be guaranteed. The most promising alternative is DCB (section 4.1.5), even though largely undocumented. If an outlier-free and full documented approach is needed, the choice has to be made between VNG (section 4.1.3) and MHC (section 4.1.4).

Signal scan interpolation: The nearest neighbor scan (section 4.2.1) is discontinuous and scattered. In contrast, bi-linear interpolation (section 4.2.2) is continuous even across cell boundaries and at the same time does not produce signal overshoots. Bi-cubic interpolation (section 4.2.3) delivers smoothest scans and even the derivative is continuous across cell boundaries but takes a 4×4 neighbourhood instead of 2×2 (bi-linear) or only one pixel (nearest neighbor) to obtain corresponding intensity values. This large scan extent affects the measurement, especially for signals close to Nyquist-frequency. Furthermore, bi-cubic interpolation occasionally is prone to produce signal overshoot. As a compromise between smoothness, no signal overshoot and medium area of influence bi-linear shall be considered as sweet spot and should be the default signal scan method.

MTF approximation: Gaussian fit (section 4.3.1) and polynomial fit (section 4.3.2) deliver functions with known derivatives in the whole range of definition even though any order of MTF derivative is rarely needed. On the other hand, both approaches use all measurements as input and if these are falsified, e.g. due to inhomogeneous target illumination or sharpening artefacts, approximated values may vary tremendously.

Main advantage of piecewise linear interpolation (section 4.3.3) is that no general mistakes are made as it occurs with Gaussian and polynomial fitting procedures. Under the premise that the resolution limit is obtained as k_{MTF10} and the piecewise linear method is unresponsive to falsified measurements at lower spatial frequencies, this shall be the default method for MTF approximation.

Siemens-star center position: Three different approaches are available to determine the correct Siemens-star center position, using a) external markers (section 4.4.1) if included in the test target b) line segments (section 4.4.2) and c) signal phase shift (section 4.4.3). It has been shown that all three approaches deliver very similar and reliable results and can be used complementary (depending on availability). A further confidence interval can be given as the distribution of the MTF itself in close vicinity of the star center and it is useful to add that value and corresponding position to the measurement report.

Normalization of contrast magnitude: Normalization with contrast level C_0 is an essential step of the approach using Siemens-stars to determine spatial resolution. This is done by calculating C_0 as the amplitude between intensities $I_{max}(f)$ and $I_{min}(f)$ for the outermost circle or by using dedicated patterns included in the test target layout. Even if the outermost circle contains pure black and white pixels the aforementioned normalization patterns can still serve as an additional control instance. It is therefore advised to include these elements in the target design.

Test target inclination: Inclination affects spatial resolution measurement in terms of finding correct extrema and subsequently obtained contrast and modulation magnitude. Having knowledge about $n \geq 4$ image points (e.g. external markers determined by an operator) and corresponding object space coordinates and scale, target inclination can be considered as a projective transformation of the scan coordinate system to avoid resampling. As both parameter sets almost never will be free of error, calculation and utilization of the projective transformation will unavoidably induce errors to the scan and the shape becomes scattered with rising manifestation for larger spatial frequencies which in turn leads to sometimes over- or underestimated extrema. Due to this fragile behavior two suggestions shall be provided. If oblique camera systems are investigated the used test targets should ideally be placed inclined on the ground to correct the oblique camera angle to a quasi-nadir setting. If this is not possible or cameras with large aperture angles are under investigation the overall measurement confidence has to be decreased and obligatory mentioned in the measurement protocol.

Motion blur: Motion smear in aerial images (mostly due to dominant 1D-forward motion) can be captured as shape of the 2D-PSF using a Siemens-star test target. With rising motion during identical exposure time frames the shape becomes more and more elliptic instead of circular. Then the quotient of semi-minor and semi-major ellipse axis can be used as measure. To determine if smear due to complex (overlapping) motions is present it is advised to obtain an initial estimate for static and dynamic conditions (and their relation). A measurement of circular shaped

2D-PSF but significantly worse spatial resolution compared to the initially calibrated dynamic estimate then indicates a complex motion during the exposure time frame.

5 Methodology Validation and Applications

The presented technique, using spokes target (namely Siemens-stars) to determine spatial resolution, must reproduce respectively confirm several (sometimes trivial) assumptions or hypothesis if implemented correctly and following all aspects described in the previous chapter. Furthermore, an approach delivering standardized and objectively determined measurements opens up a variety of application options in the field of remote sensing sensors. Therefore this chapter is thematically divided into two parts.

The first part describes and validates aforementioned assumptions and hypothesis, i.e., exposure time independence (section 5.1), number of star segments independence (section 5.2), coherence between sinusoidal- and square wave Siemens-stars (section 5.3), comparison to Slanted-edge approach and bar chart method (section 5.4), measurability of applied predefined model parameters (section 5.5) and measurement repeatability (section 5.6).

The second part describes the setup and delivered results for various applications, i.e., a camera resolution benchmark (section 5.7), an in-field sensor validation (section 5.8) and an image field MTF representation (section 5.9).

5.1 Exposure Time Line

5.1.1 Expected Measurement

Determination of modulation transfer function for identical optical sensor systems is exposure time independent, because the ratio between intensity maximum $I_{max}(k)$ and intensity minimum ($I_{min}(k)$) raises linearly with increasing exposure time, depending on the amount of collected photons and would reach the unresolved common gray level at the very same spatial frequency. Furthermore, the normalization to contrast level C_0 (see equation 3.5) at spatial frequency equal to 0 (infinite radius) ensures values for MTF $\tilde{H}(k)$ to be in range between 0 and 1.

Assuming applied series subtraction according to Coltman [66] or sinusoidal signals equation 3.5 can be rewritten as:

$$\tilde{H}(k) = \frac{I_{max}(k) - I_{min}(k)}{I_{max}(k) + I_{min}(k)} * \frac{1}{C_0} \tag{5.1}$$

Normalization with C_0 would only be obsolete if $I_{min}(0)$ is equal to zero but this is almost never true, because the used materials for resolving power targets (especially parts for the black patterns) do not fully absorb all photons but reflect a significant portion and make the use of a normalization factor mandatory.

Additionally, the signal to noise ratio will decrease for lower exposure time frames and will gain influence, assuming all other image acquisition parameter can be kept constant.

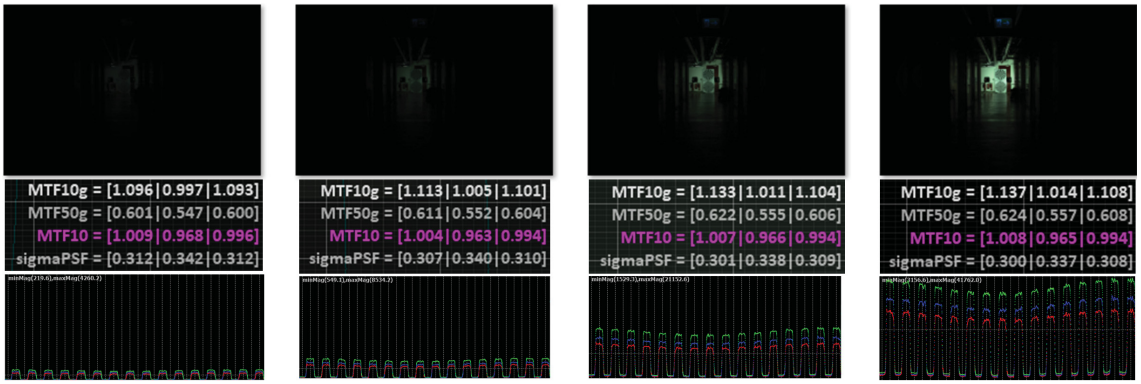


Figure 5.1. (top row) images at different exposure times (mid row) determined MTF values (bottom row) signal scans for different exposure times

5.1.2 Validation

Aforementioned theory can be proven empirically by determining spatial resolution for different exposure time frames while keeping all other acquisition parameters constant. This can be achieved using a sufficiently large basement hall (alternatively a shaft, tunnel or huge chimney) and illuminating the target with controlled and diffuse light (see Figure 5.1).

The experiment has been conducted for five different sensor-lens combinations. For every camera four images at different exposure times have been acquired (low, medium low, medium high and high exposure). The results, showing the maximum difference to the mean value of all four different exposure times, are given in Table 9. A mean deviation of 2.53% for all channels (red, green, blue) is very low and confirms the introduced hypothesis that determination of modulation transfer function for identical optical sensor systems is exposure time independent.

5.2 Number of Star Segments

5.2.1 Expected Measurement

Similar to independence of exposure time, also the number of Siemens-star segments should not have any influence on determination of modulation transfer function if all other parameters are kept constant (identical sensor-lens combination, spokes target material, illumination, exterior orientation). But still one aspect remains influential, i.e., the positional relation between scan pattern and scenery (see Figure 5.2). It can be seen that spatial resolution appears clearly different even though all aforementioned parameters are constant. This difference occurs due to aliasing effects and depends on the sub-pixel relation between camera and scenery. If, for example, one Siemens-star is positioned in favour of avoiding most aliasing artefacts and another (identical) star is located at a non-integer position offset imaged and determined spatial resolution will be slightly different. This positioning, however, is

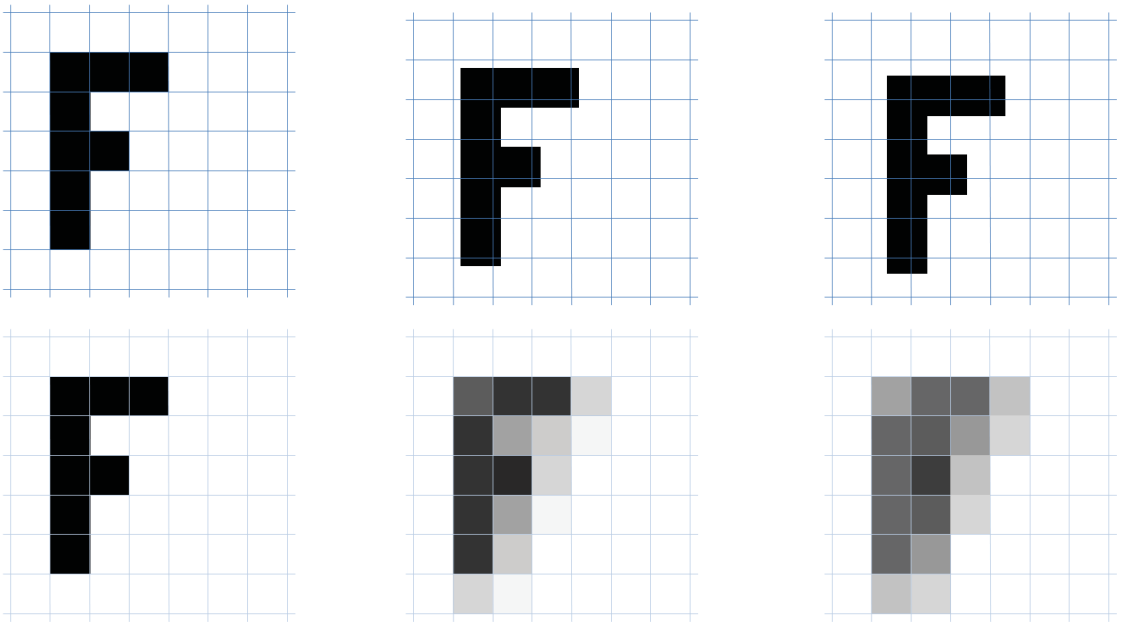


Figure 5.2. Spatial resolution dependent on positional relation between scan pattern and scenery (top) group of bars in relation to scan pattern (bottom) rendered results

a random factor in general and can not be foreseen and is one aspect why the overall measurement accuracy can not reach 100%.

Nevertheless, using Siemens-stars with more segments (and thus more support values) is advantageous because it increases the chances that aliasing effects can be averaged alongside a circle scan and subsequent resolution determination. At the same time more segments have the disadvantage of increasing overall target size necessary to measure spatial resolution (see section 4.5).

5.2.2 Validation

Analogical to independence of exposure time, above-named hypothesis can be proven empirically by determining spatial resolution for different Siemens-stars (different number of segments) while keeping all other acquisition parameters constant.

The experiment has been conducted for twenty images of six different sensor-lens combinations each having imaged one Siemens-star with 16 black-white segments and another with 36 black-white segments in the center of the image plane. The results, showing the difference between measurements of both stars (min/max) are given in Table 10. A mean deviation of 2.43% for all channels (red, green, blue) can be considered low and confirms the introduced hypothesis that determination of modulation transfer function for identical optical sensor systems does not dependent on the number of Siemens-star segments.

5.3 Sinusoidal- and Square Wave Siemens-stars

5.3.1 Expected Measurement

In 1954, John Coltman postulated the theory on how to obtain MTF values using square wave targets [66]. He described the approach as follows:

"The square wave response factor $r(n)$ may be obtained experimentally by using a black and white bar pattern and measuring the maxima and minima of the image brightness for various values of n. We wish to find $R(n)$, the sine wave response, knowing $r(n)$. This series may be solved for R(n) by successively subtracting series for r(kn)/k, using values of k chosen by inspection so as to eliminate higher and higher terms in R(kn). There results the series:"

$$R\left(n\right) = \frac{\pi}{4}\left[r\left(n\right) + \frac{r\left(3n\right)}{3} - \frac{r\left(5n\right)}{5} + \ldots\right] \tag{5.2}$$

This implies that the contrast transfer function obtained using sinusoidal Siemens-stars must be identical in comparison to the modulation transfer function of a Siemens-star with pure black and white segments.

5.3.2 Validation

The implication and thus Coltman's hypothesis can be proven empirically by an experiment. Therefore, five different images have been used where a sinusoidal Siemens-star and a square wave Siemens-star are placed and imaged close to one another (see example Figure 5.3). For the sinusoidal star the contrast transfer function is obtained and subsequently compared to modulation transfer function obtained of the black and white square.

The results, showing the difference between measurements of both stars (min/max) are given in Table 11. A mean deviation of 2.90% for all channels (red, green, blue) is very low and confirms both Coltmans's hypothesis and the implication that CTF of sinusoidal Siemens-stars and MTF of square wave Siemens-stars are identical.

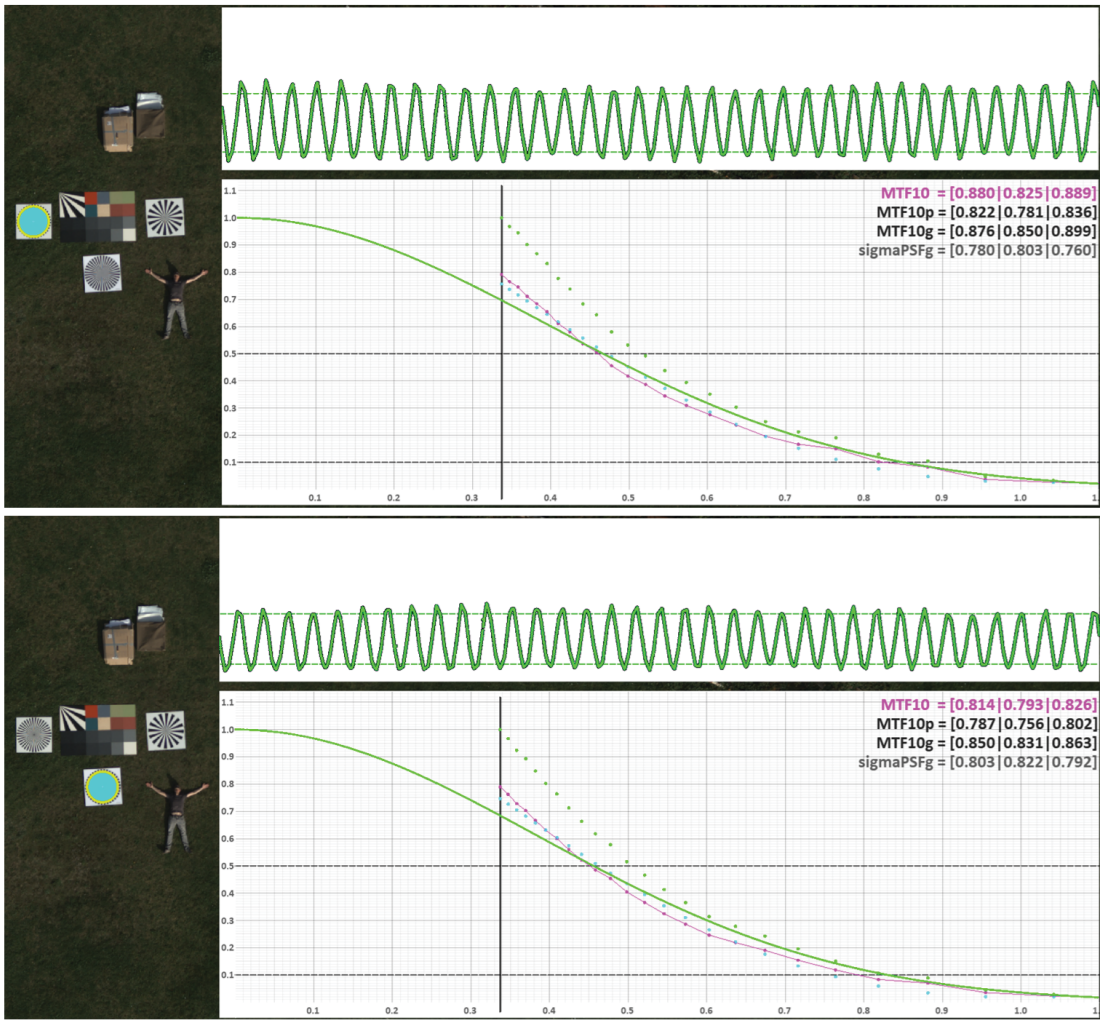


Figure 5.3. Square wave MTF measurement (top) in comparison to sinusoidal wave CTF measurement (bottom)

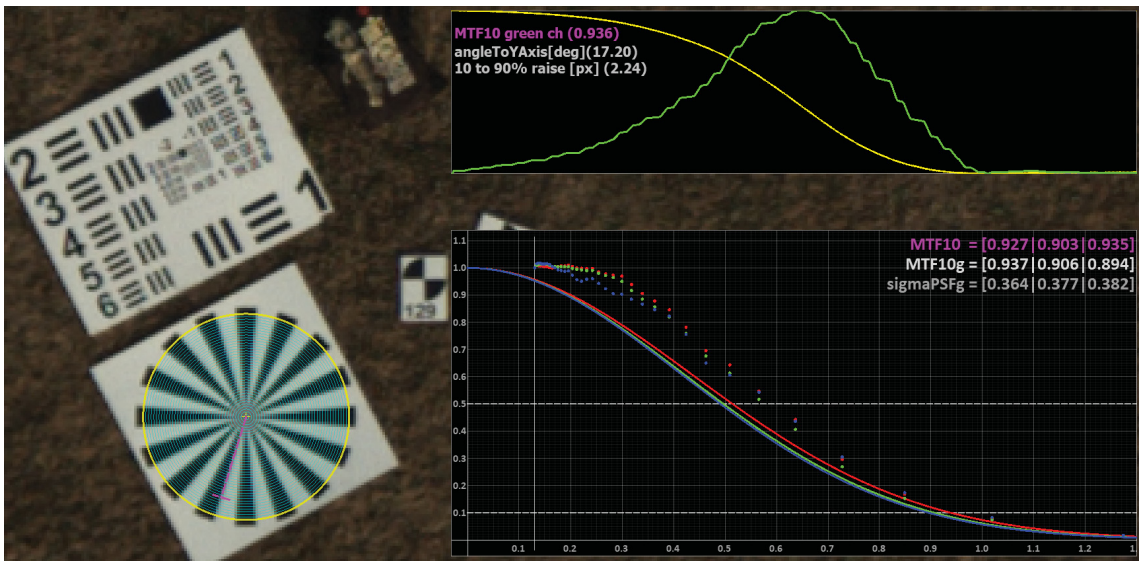


Figure 5.4. Simultaneous determination of ground resolved distance (GRD) operator driven for USAF51 (top left), Slanted-edge from ESF & LSF (top right) and Siemens-star (bottom right)

5.4 Comparison to Slanted-Edge Approach and Bar Chart Method

5.4.1 Expected Measurement

Given the variety of approaches and techniques it is consistently necessary to compare their respective measurement results and answer the question if used techniques do or do not perform equivalently and what are reasons for particular observations.

Therefore, all described approaches (bar targets, Slanted-edge and Siemens-star) have been applied simultaneously for identical images and image regions (example Figure 5.4). Used image quality parameter is ground resolved distance (GRD in cm). For USAF51 bar chart GRD is calculated according to equation (3.1) and the number of observers is $n \geq 10$. Modulation transfer function for Slanted-edge (derived from ESF and LSF) is calculated by using one edge of the Siemens-star (see Figure 5.4, magenta line, lower left). Reciprocal of MTF10 for both, Slanted-edge approach and Siemens-star technique, is the approximation for size of the smallest line per pixel if the unit for MTF is line per pixel (or pixel/pixel) which is the case for both implementations. Multiplying reciprocal of MTF10-values from Slanted-edge and Siemens-star measurement with calculated ground sample distance (GSD, equation 3.2) delivers GRD for both algorithms.

5.4.2 Validation

Ten images, showing bar chart and Siemens-Star simultaneously at GSDs between 1.24cm and 1.27cm, have been taken by a sensor-lens combination mounted on a

fix-wing drone. As mentioned in the previous section, GRD has been calculated for every image and every approach/technique. Related results are given in Table 12).

There, the fourth column (Δ SLE-Star) shows the absolute difference between Slanted-edge and Siemens-star in percent. Except for one outlier, both techniques measure very similar at an overall mean difference of 3.3%. Values obtained by independent human observers and USAF51 tend to be more static compared to the other methods. This effect is most likely caused by huge resolution steps between groups of bars. Rearranging the target, including more groups with finer descent, may weaken the effect. Even when comparing all three approaches at once (Δ Min-Max) the overall mean difference of 7.1% still can be considered very low.

5.5 Model-Based MTF and PSF

5.5.1 Expected Measurement

A conclusive validation of Slanted-edge (section 3.2) and Siemens-Star technique (section 3.3) is to apply predefined modulation (MTF) or spread parameters (PSF) to an ideal representation of resolving patterns (see Figure 5.5). This can be done in both domains. In image-domain it can be done by forming a convolution of mathematical-ideal image-intensity values (I), a Gaussian-shape model PSF (H_m) and a mathematical-ideal sensor PSF (H_s). Simulated PSF (H_{sim}) then can be formulated as follows:

$$H_{sim}(r) = I(r) * H_m(r) * H_s(r) \tag{5.3}$$

In frequency-domain calculation gets simpler, only the product of image spectrum ($\tilde{I}$) with a predefined model-based MTF ($\tilde{H}_M$) and (mathematical-ideal) sensor MTF ($\tilde{H}_S$) has to be calculated. Therefore, simulated-image MTF ($\tilde{H}_{Sim}$) can be formulated as follows:

$$\tilde{H}_{Sim}(k) = \tilde{I}(k) \cdot \tilde{H}_M(k) \cdot \tilde{H}_S(k) \tag{5.4}$$

The derived hypothesis is, if both algorithms (Slanted-Edge and Siemens-Star) described in chapter 3 provide measurements of absolute value then model-MTF ($\tilde{H}_M$) respectively model-PSF (H_m) must directly be confirmed by measurement of simulated-image MTF ($\tilde{H}_{Sim}$) respectively PSF (H_{sim}).

Mathematical-ideal aliasing-MTF $\tilde{H}_S(k)$ with $k \in \mathbb{R}$ is characterized as being equal to one for all frequencies up to frequency belonging to the pixel-size and zero for all frequencies above that threshold (see Figure 5.5, dotted line). However, when an ideal pattern is rendered to a pixel grid the resulting (Nyquist-limited) aliasing-PSF and aliasing-MTF unavoidably will differ from ideal shape. An example of (close to) ideal aliasing-MTF can be seen in Figure 5.5 (dashed-dotted line) with Nyquist-limit 1.0 line per pixel.

As a result, obtained MTF values ($\tilde{H}_{Sim}$) reflect the product of (close to) ideal

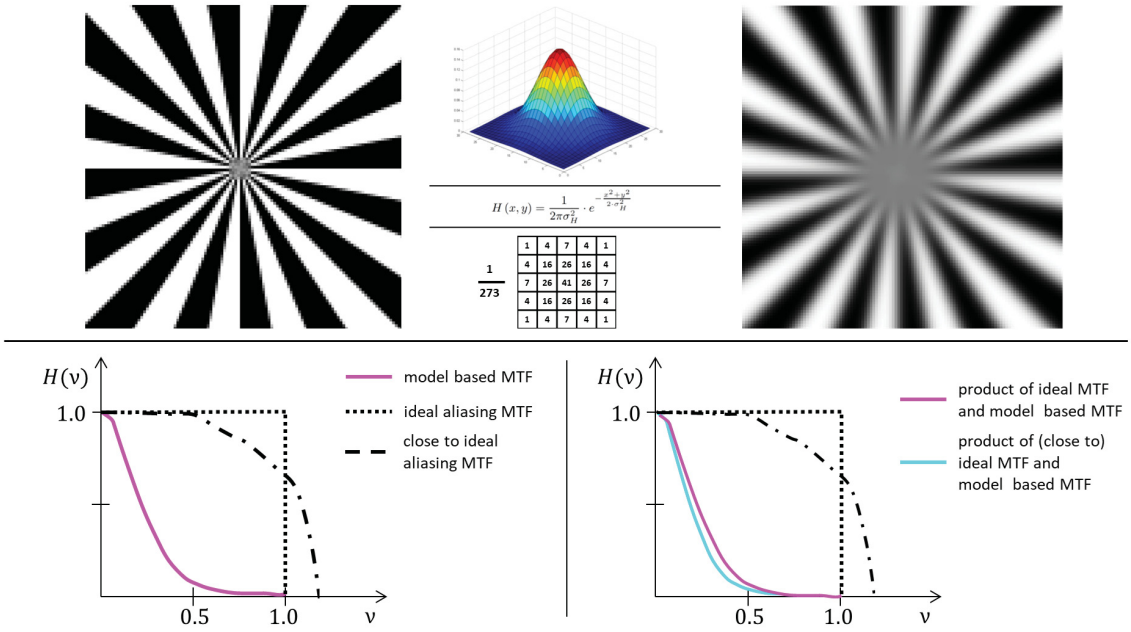


Figure 5.5. Original image (upper left) continuous and discrete Gaussian PSF convolution kernel (upper mid) and convolution result (upper right), ideal aliasing MTF, close to ideal aliasing MTF and model-based MTF (lower left), related products in frequency domain (lower right)

aliasing-MTF ($\tilde{H}_S$) and model-MTF (H_M) and therefore are expected to be smaller than the product of ideal aliasing-MTF ($\tilde{H}_S(k) = 1,\ k \in \mathbb{R}$) and model-MTF ($H_M$) (see Figure 5.5, magenta and cyan line).

Considering that PSF and MTF are directly linked by (inverse) Fourier transformation (equation 3.8), it can be assumed that for increasing values σ_m (H_s) respectively for decreasing values σ_M ($\tilde{H}_S$) simulated images and corresponding measured quality parameter σ_{SLE} and σ_{Star} of H_{sim} will be continuously less affected by the difference of ideal and (close to) ideal aliasing-PSF or MTF.

5.5.2 Validation

Aforementioned assumption can be verified by an experiment as a conclusive validation of Slanted-Edge and Siemens-Star technique applying predefined modulation (MTF) or spread parameters (PSF). Then, the used model parameters σ_m must be reproduced/measured by both methods during measurement (σ_{SLE} and σ_{Star} of H_{sim}). For this reason, an image showing a Siemens-Star including (close to) ideal aliasing-PSF (H_s) has been convolved with different σ_m starting at 0.500 and rising to 1.750. Subsequently, σ of H_{sim} has been calculated using both approaches, Slanted-edge and Siemens-Star. Obtained results can be found in Table 13.

Values in column $\Delta(\sigma_m, \sigma_{SLE})$ show the difference between predefined and applied model parameter σ_m and measured parameter σ_{SLE} in absolute percentage [%]. Val-

ues of column $\Delta(\sigma_m, \sigma_{Star})$ provide results for difference between σ_m and measured σ_{Star}. Similar to the comparison of both techniques in preceding section 5.4 column $\Delta(\sigma_{SLE}, \sigma_{Star})$ reflects the absolute difference between σ_{SLE} and σ_{Star} in absolute percentage [%].

Two observations can be emphasized.

First, the difference between Slanted-Edge and Siemens-Star technique again is very small. In contrast to experiment described in section 5.4 overall mean difference of 1.6% here is even smaller and measurements deliver no outliers.

Second observation regards constructed hypothesis described in previous subsection 5.5.1: "... it can be assumed that for increasing values σ_m (PSF) ... simulated images and corresponding quality parameter ... will be continuously less affected by the difference of ideal and (close to) ideal aliasing-PSF". Columns $\Delta(\sigma_m, \sigma_{SLE})$ and $\Delta(\sigma_m, \sigma_{Star})$ indicate that this hypothesis is true. With rising σ_m the absolute difference of both methods tend to approach zero.

5.6 Measurement Repeatability

5.6.1 Expected Outcome

The last validation issue concerns measurement repeatability and user / operator independence.

As already stated in section 3.1.1, the use of bar targets (e.g. USAF51 target) and the operator driven estimation of spatial resolution is very common but contains subjective influence, even though the mean observation of many operators is used to diminish the impact.

Aim of this verification is to proof that measurements, obtained by the implemented software tool as support for the norm DIN 18740-8 "Photogrammetric products – Part 8: Requirements for image quality (quality of optical remote sensing data)" and using Siemens-star targets, are reliable and user independent. If the program delivers measurements free of subjective influence, then all related values must be similar (ideally identical) for every processed image according to the procedure described in section 3.3.

5.6.2 Validation

Therefore a statistical relevant number of operators ($n \geq 10$) conducted measurements using the aforementioned software tool to determine spatial resolution for a statistical relevant number of images ($i \geq 10$). Results are given in table 14 (Appendix).

It is obvious that the workflow and underlying implementation delivers very reliable results. Although, operator accuracy for Siemens-star center determination has a mean difference of 5.2pixel the refined position is the exact same for all independent measurements. Very similar observations can be made for spatial resolution parame-

ters itself. The mean difference for all (MTF) measurements of all operators is 0.11% and thus very low.

Presented results strongly indicate that the software tool as official part of upcoming norm DIN 18740-8 "Photogrammetric products – Part 8: Requirements for image quality (quality of optical remote sensing data)" delivers highly reliable measurement parameters which are furthermore free of subjective influence.

5.7 Camera Resolution Benchmark

This section and all following sections in this chapter present various applications considering the underlying theory of resolution testing for different sensor-lens combinations.

By this time there is a growing variety of camera systems that either claim to be suited for remote sensing (mostly UAV) applications or could be considered suitable because of their specifications (e.g. in terms of sensor size, trigger event control, overall weight, interface accessibility and acquisition costs). These systems are often being grouped as a) systems specifically designed for remote sensing (mostly UAV) purposes b) large format cameras, c) system (or bridge-) cameras and d) single-lens reflex cameras.

The camera systems compared in this thesis are given in Table 1 where Sigma's DP1 is listed in the category *Other*. The decision was made due to the fact that this camera comes with a Foveon-sensor [90] where each pixel element detects color information instead interpolating neighbourhood pixel of a Bayer-pattern arrangement.

5.7.1 Benchmark Layout

The experiment, which is also appropriate as a benchmark procedure, is motivated and arranged as follows.

Aim is to determine and compare spatial resolution for different sensor-lens combinations. In order to guarantee repeatable measurements with identical controlled light conditions and to prevent extraneous light, a sufficiently large basement hall has been chosen (see Figure 5.6 top). For every camera system the related distance to the designated test pattern (see Figure 5.6 bottom) has been calculated to ensure identical nominal ground sampling distance (GSD - according to focal length and pixel-size of the sensor). The GSD in this benchmark has been set to 5mm to address the designated fields of application (e.g. inspection / monitoring) including their resolution requirements. Usually, resolving power is changing across the field of view. In order to analyse this effect multiple images have been taken to have the resolution target imaged at different locations in image space (e.g. image center – image half field – image corner). All images have been converted from their raw format to usable tiff format using the same demosaicing algorithm. In this case, the DCB method (see Sub-section 4.1.5) with its open source (C++) implementation of <libraw> and the primordially implementation for DLR's MACS Micro versions <mipps>has been

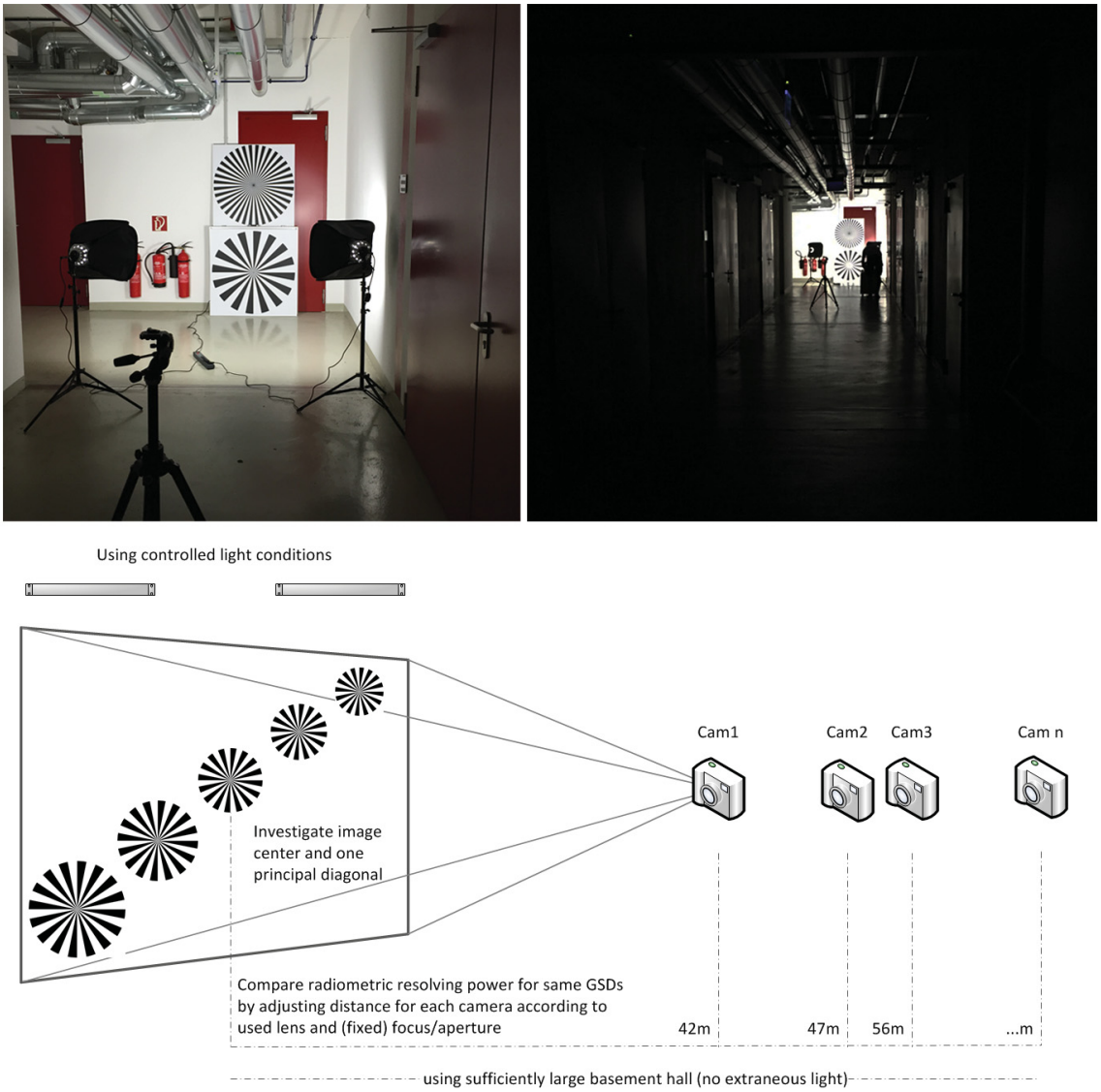


Figure 5.6. Controlled light conditions (top left), no extraneous light (top right). Illustration of the experiment setup (bottom)

Camera Class	Purpose built / Proprietary										Large Format		System / Bridge Cameras				DSLR		Other	
Camera	DJI Inspire 1 Pro 12mm lens		DJI Inspire 1 Pro 15mm lens		DJI Phantom 3 Advanced		MACS Micro 24mm lens		MACS Micro 50mm lens		PhaseOne IXU1000 55mm lens		Ricoh GXR Mount A12 21mm lens		Sony Alpha 7R 35mm lens		Canon EOS 5DS R 24mm lens		Sigma DP1	
Lens	Olympus M. Zuiko 12mm f2.0		DJI MFT ASPH 15mm f1.7		Integrated lens 3.6mm f2.8		Schneider Kreuznach Apo-Xenoplan		Leica APO Summicron 50/2.0		Rodenstock 50mm f/5.6		Zeiss Biogon 21/2.8		Voigtländer Skopar 35/2.5		Canon EF 24 f/1.4L II USM		Integrated lens	
MPx	16		16		12		12		16		100		12		36		52		4,7*	
Approx. weight (incl. lens)	545	g	545	g	188	g	320	g	720	g	930	g	916	g	591	g	1626	g	277	g
Fixed aperture	4.0	f	4.0	f	2.8	f	4.0	f	2.8	f	4.0	f	7.1	f	5.6	f	5.6	f	5.6	f
Pixel pitch	3,75	µm	3,75	µm	1,55	µm	4,7	µm	7,4	µm	4,6	µm	5,5	µm	4,87	µm	4,14	µm	7,8	µm
Focal length	12	mm	15	mm	3,61	mm	24	mm	50	mm	50	mm	21	mm	35	mm	24	mm	16,6	mm
Pixel count in X	4608	px	4608	px	4072	px	4000	px	4864	px	11608	px	4288	px	7360	px	8688	px	2640	px
Pixel count in Y	3456	px	3456	px	3044	px	3000	px	3232	px	8708	px	2848	px	4912	px	5792	px	1760	px
FOV diagonal	84,0	°	71,5	°	95,0	°	52,2	°	46,7	°	67,4	°	68,0	°	67,2	°	84,0	°	73,4	°
FOV in X	71,5	°	59,9	°	82,3	°	42,8	°	39,6	°	56,2	°	58,6	°	57,9	°	73,7	°	63,6	°
FOV in Y	56,7	°	46,7	°	66,3	°	32,7	°	26,9	°	43,7	°	40,9	°	40,5	°	53,1	°	44,9	°
Image Acquisition Parameter																				
Distance	16,0	m	20,0	m	11,7	m	25,6	m	33,8	m	54,2	m	19,1	m	36,0	m	29,0	m	10,7	m
Native ground resolution (image center)	5,0	mm	5,0	mm	5,0	mm	5,0	mm	5,0	mm	5,0	mm	5,0	mm	5,0	mm	5,0	mm	5,0	mm

Table 1. Specifications of the compared camera systems (* SigmaDP1's Foveon-chip with *true* 4.7Mpx has no Bayer-pattern, therefore no demosaicing necessary)

chosen. For every converted image the PSF and corresponding σ (see section 3.3.1) has been calculated.

This benchmark setup should guarantee the genuine system response from object space to sensor. The results are presented in Figure 5.7.

5.7.2 Results

The (effective) resolving power of a specific sensor-optic (camera-lens) combination decreases as expected from image-center via image-half-field to the image-corner. This deterioration of resolving power is caused by radially symmetric lens distortion and vignetting- effects as clearly can be seen looking at the trend of Sony's Alpha 7R (with Voigtländer Skopar 35/2.5 lens). Whilst the resolving power in center area is close to top-rated systems it deteriorates extraordinarily to image border.

Imaging performance of all DJI systems is fairly homogenous and the variation in resolving power in comparison to the aforementioned sensor-optic combination is quite low. The MACS Micro prototype system (especially the 16 megapixel version) shows top-rated results which indicates a connection between pixel-size (and thus SNR) and resolving power since this sensor's pixel pitch is the largest (7.4 μm) compared to other systems with Bayer-pattern. The results for resolving power of Sigma's DP1 support the assumption since its pixel pitch is the most largest (7.8 μm).

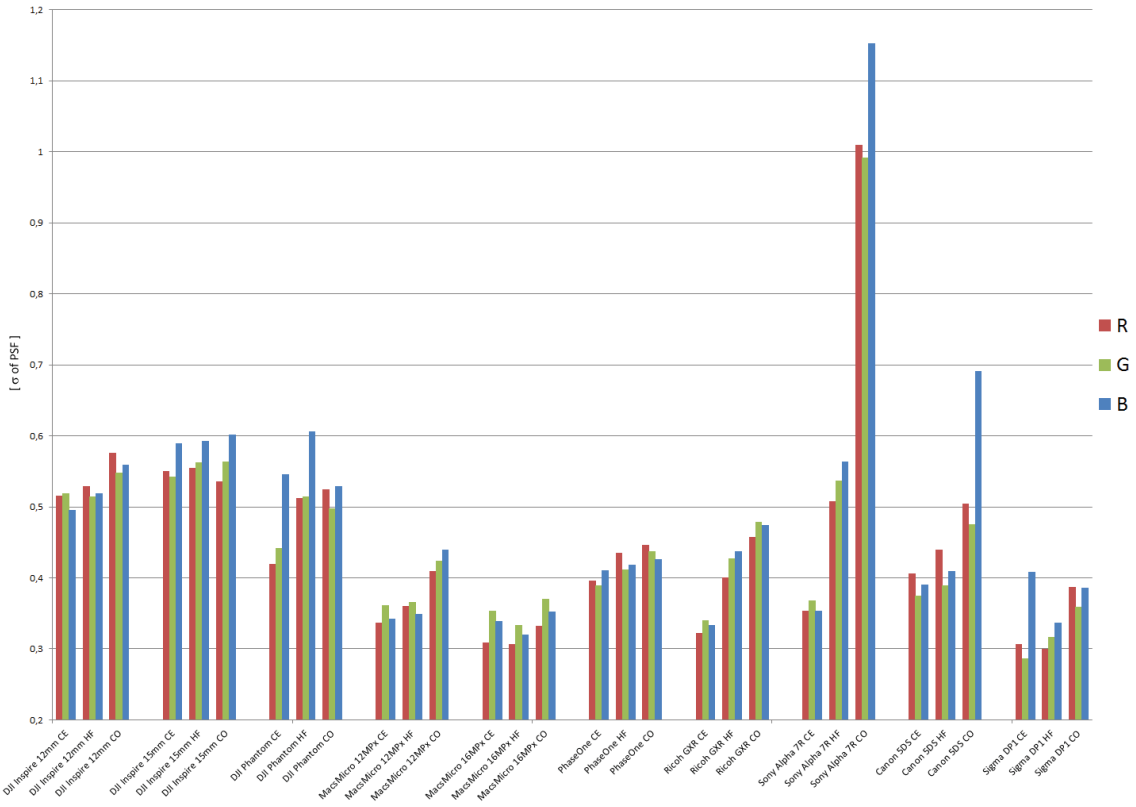


Figure 5.7. Results for σ of PSF in pixel (CE center, HF half field, CO corner)

Furthermore the Foveon-chip outperforms all other systems and this in turn highly indicates the large impact of demosaicing (Sigma's DP1 without a Bayer-pattern and thus demosaicing is not necessary).

Ricoh's GXR offers a consistent performance and is close to the MACS Micro (12MP) and Phase One's IXU 1000. Especially the image center resolves close to top-rated value.

The Phase One IXU 1000 with its large sensor format and 100 megapixels shows very low variation in resolving power over the complete image space although it is slightly behind in terms of overall spatial resolution compared to MACS Micro.

Canon's EOS 5DS R as a representative of the DSLR group is equal to PhaseOne's IXU but its resolving power diminishes perceptively along the image diagonal.

As it can be seen the blue channel occasionally is determined significantly worse compared to green channel and especially red channel. The Bayer-pattern arrangement consists of twice the number of pixel for green compared to red respectively blue. Therefore one would expect slightly better results for the green channel but almost equal results for red and blue. This issue is not finally solved and will be investigated in future work considering the presumptions if this problem is either caused by chromatic aberrations and / or color temperature of the used light source.

5.8 In-field Sensor Validation

5.8.1 Application

Spatial resolution parameter of a camera system in motion is expected to be lower than under (static) laboratory conditions. Main reason is the sensor movement (6 DOF) through 3D-space but with additional influence of the carrier-unit's (micro-)vibrations under operating conditions (see section 4.7). Therefore a true expected value for achievable spatial resolution of a specific sensor-lens combination can be obtained by determine and compare spatial resolution in both cases.

Prepared and correctly executed flight plans with image overlaps of 80% in track and 60% across track ensure the spatial resolution test target (Siemens-star) to be in at least 15 images. Simultaneously, the test target is imaged at different image locations and similar to the benchmark procedure under static conditions (see section 5.7) images containing the Siemens-star at center-, half-field and corner-position can be selected for measurement and comparison.

5.8.2 Results

As explained in preceding subsection, spatial resolution parameter for static (laboratory) conditions are expected to be better than under kinematic (operating) conditions due to motion caused by carrier-movement and -vibrations. Further, spatial resolution is expected to be better when imaged in center than in image corners.

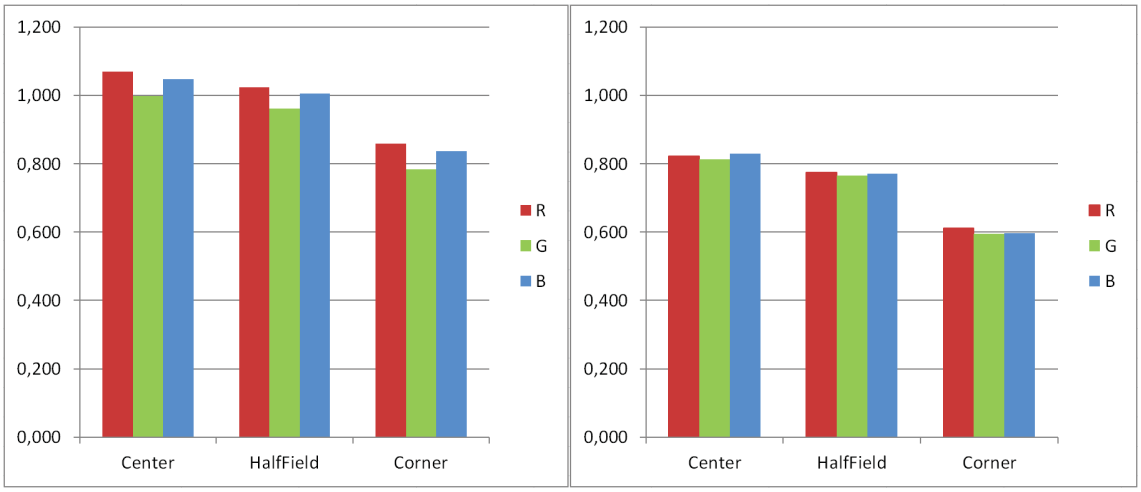


Figure 5.8. Results for resolving power measurements [line/pixel] under static (left) and kinematic (right) conditions.

To investigate both issues, images under static conditions containing the Siemens-star at center-, half-field and corner-position have been selected for measurement. The same selection has been made for kinematic conditions. For every image the value for MTF at 10% contrast level (resolution limit, see section 2.1.2) has been calculated and serves as measure to conduct the analysis. Results for both aspects are given in Figure 5.8.

As expected, resolution limit (MTF10) declines from image-center via image-half-field to image-corner for both static and kinematic conditions. This deterioration is mainly caused by chromatic and spherical aberrations induced by the lens-system. Furthermore, the mean difference between operating and static conditions is 23.6% (center 20.9%, half-field 22.7%, corner 27.2%). Considering very high image performance under laboratory conditions (around 1.00 line/pixel and thus near Nyquist-limit) a reduction of 23% of resolution under operating conditions still can be considered satisfactory.

5.9 Image-field MTF

5.9.1 Application

While spatial resolution parameters for static (laboratory) and kinematic (operating) conditions can be determined across the image diagonal (see previous sections) an even more complete and convenient way is to measure and analyse these parameters across the entire image-field (see Figure 5.9 (bottom)).

Many flights, especially for in-situ calibration and high precision surface reconstruction are not only conducted with aforementioned image overlaps of 80% in track and 60% across track but as a cross flight pattern to strengthen the block geometry for subsequent bundle block adjustment [91, 92, 93]. Then, the Siemens-star is imaged

at least in 30 different locations which all serve as sample points for the image-field MTF representation.

This kind of depiction has the advantage that not only the grade of spatial resolution decrease from image center to image border can be considered but also possible asymmetrical behaviour of the sensor-lens combination under investigation.

5.9.2 Results

The algorithm to deliver spatial resolution as an image-field representation is implemented as follows.

The image coordinate system is divided into a predefined number squares, e.g. 8 $\times$ 5, see Figure 5.9 (top). For every node (joint corners of adjacent squares) a value (MTF10) is calculated either as linear interpolation or bi-linear interpolation of opposing values depending on availability of sample points. This step is being performed repeatedly (while-loop) until every node has got an assigned value. As a side note, the four image corner nodes can not be interpolated, therefore, the closest value (pixel-distance) is assigned to that position.

Having established a regular grid, the final step consists of bi-cubic interpolation across the whole wire-frame including subsequent colorization according to minimum and maximum MTF10-values.

The final result is given in Figure 5.9 (bottom). As expected, the best resolved image part is the center and deteriorates to image corners. The camera system under investigation in this case was DLR's MACS-Micro with 16 megapixel and a Leica APO Summicron lens with an aperture angle of 46.7° (diagonal). Advantage of this rather narrow aperture angle in combination with the APO Summicron lens is the very homogenous resolution across the complete image field. Except for the image corners, where spatial resolution deteriorates significantly, MTF10 values are between 0.750 and 0.850 line/pixel.

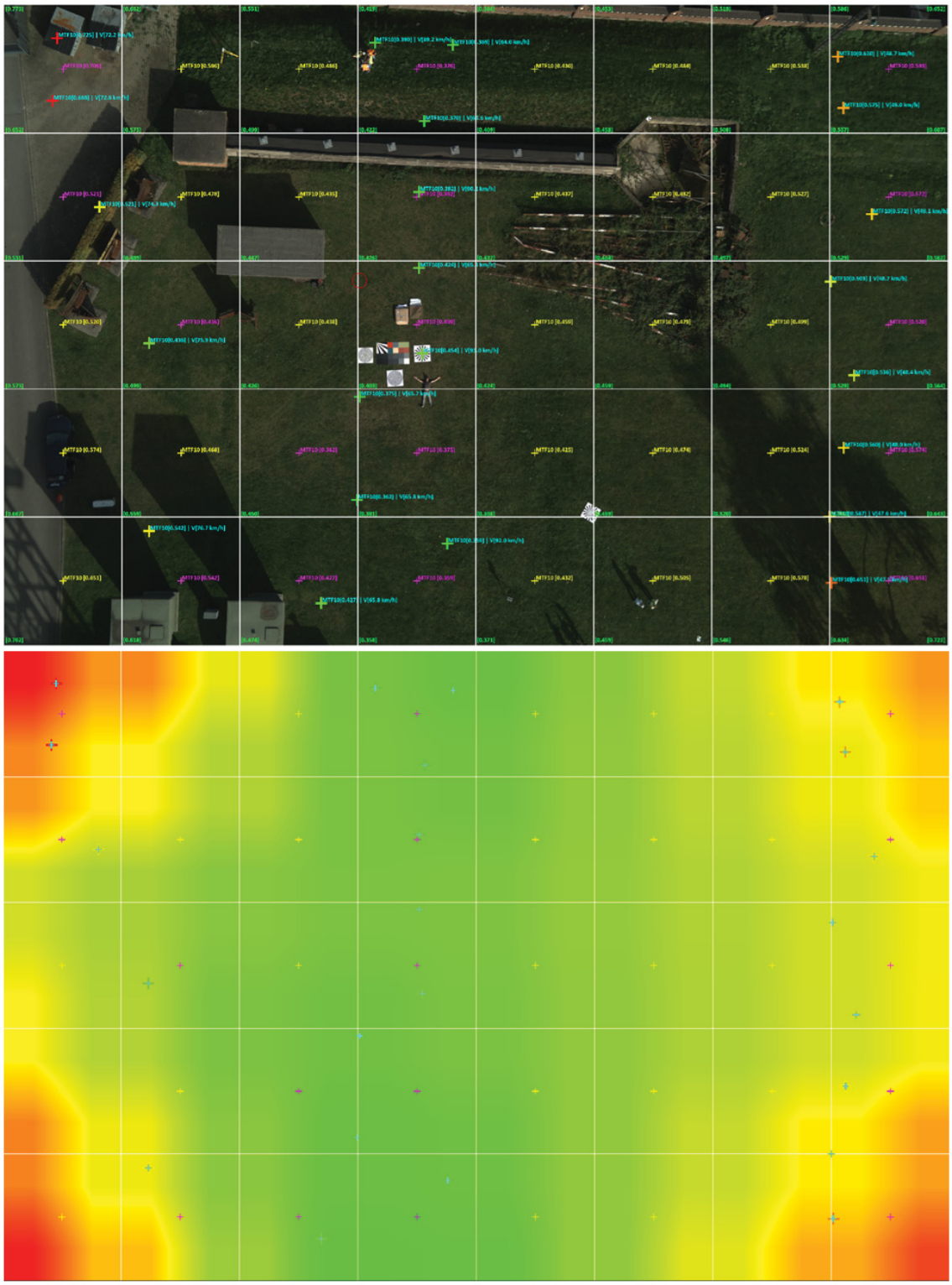


Figure 5.9. Image-field MTF, (top) example image with sampling points and interpolation grid, (bottom) interpolated image-field MTF

6 Resolution Improvement

Improvement of image resolution (in this thesis with focus on remote sensing images) can be achieved through various strategies.

A rather simple approach is edge sharpening to diminish aliasing effects. One general procedure is to use unsharp masking. There, a Gaussian smoothing filter is applied and subsequently subtracted (using weights to remain homogeneous neighboured values constant) from the original image. Another approach is to use sharpening kernels and subsequently convolute the entire image with that matrix. But also more sophisticated sharpening approaches are available where the initial edge is determined and sharpening is applied perpendicular to that direction [94, 95, 96].

On the one hand, some software packages of remote sensing systems manufacturers provide solutions to enhance images and their edges (e.g. Phase One's Capture One Suite [97]).

On the other hand, some providers have integrated edge sharpening techniques in their image acquisition pipeline where resolution improvement is implemented hardware accelerated and applied directly while the image is taken and stored (e.g. DJI's Phantom 4 and Mavic provide these techniques [98]). However, it still remains to be proven that this kind of sharpening enhances different subsequent tool chains, such as bundle block adjustment and dense image matching.

A further way of improving image resolution is to enlarge a given image while simultaneously applying conventional interpolation algorithms (e.g. bilinear, bicubic or different resampling methods) [99, 100]. These methods are called interpolation-based single image super-resolution and the use of deep learning algorithms gets more and more common [101, 102, 103].

Another way (and object of this thesis) is multi-view super-resolution. There, a certain number of low-resolved images are used to generate one or more higher (super-) resolved image(s). The following sections describe the pivotal idea of multi-image super-resolution, two state of the art approaches to enhance image resolution and a new technique which combines both described algorithms to obtain super-resolved images while making use of precise knowledge about the camera system.

The ultimate grade of improvement will be determined using the software tool described in preceding chapters 3 - 5 by comparing MTF10 values of low-resolved and high-resolved images.

6.1 Essence of Multi-Image Super-Resolution

6.1.1 Optical Flow

To construct one or more super-resolved image(s) from several low-resolved input images it is indispensable to have knowledge about their related orientation to one another. In remote sensing applications these values are obtained as exterior

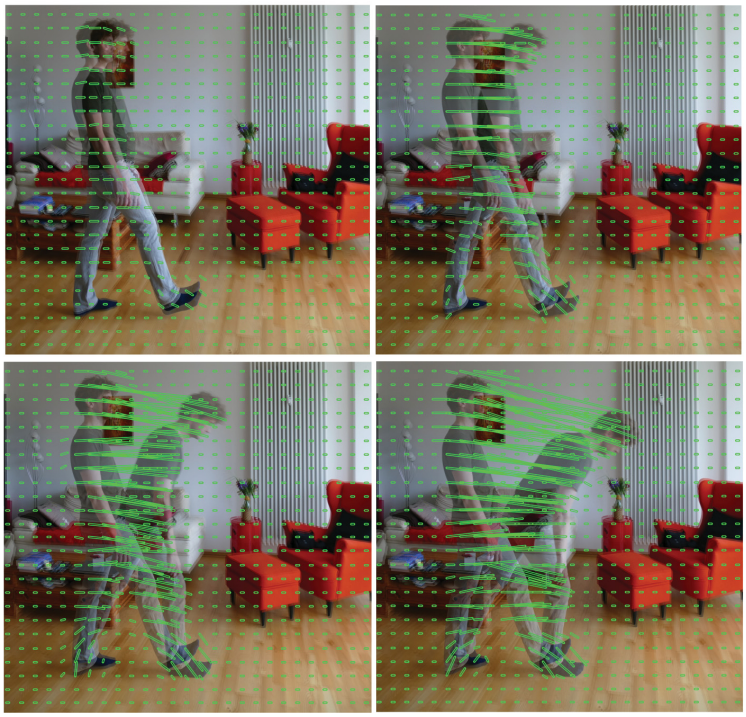


Figure 6.1. Example of an optical flow field

orientation parameters (e.g. using GNSS/INS systems) and interior orientation parameters (given by sensor and lens manufacturers) and often refined during the bundle block adjustment step (see section 2.2). However, most super-resolution approaches are applied to imagery of video streams or very high overlapping scenes. There, it is common to obtain relative orientation as a vector field (see Figure 6.1) for each pixel from one image frame to the succeeding frame as outcome of a procedure named optical flow [104].

This method starts with the assumption that image intensities of one frame $I(\vec{x}, t)$ at pixel position $\vec{x} = (x, y)^T$ at time t can be expressed as image intensities of the following frame taken at time $t + 1$ and an added 2D-velocity $\vec{u} = (u_1, u_2)^T$ [105]:

$$I(\vec{x}, t) = I(\vec{x} + \vec{u}, t + 1) \tag{6.1}$$

Further assumptions are that "the intensity structures of local time-varying image regions are approximately constant under motion for at least a short duration" [106] and that the relative movement between two frames is small. Then, the image constraint can be expanded with Taylor series where the higher order terms are most often neglected:

$$I(\vec{x} + \vec{u}, t + 1) = I(\vec{x}, t) + \vec{u} \cdot \nabla I(\vec{x}, t) + I_t(\vec{x}, t) + \cdots \tag{6.2}$$

"where $\nabla I = (Ix, Iy)$ and I_t denote spatial and temporal partial derivatives [...]. Of course, one cannot recover $\vec{u}$ from one gradient constraint since" equation 6.2 contains "two unknowns, u_1 and u_2. [...] One common way to further constrain $\vec{u}$ is to use gradient constraints from nearby pixels, assuming they share the same 2D velocity. With many constraints there may be no velocity that simultaneously satisfies them all, so instead we find the velocity that minimizes the constraint errors. The least-squares [...] estimator minimizes the squared errors" [107]:

$$E(\vec{u}) = \sum_{\vec{x}} g(\vec{x}) \left[\vec{u} \cdot \nabla I(\vec{x}, t) + I_t(\vec{x}, t)\right]^2 \tag{6.3}$$

"where $g(\vec{x})$ is a weighting function that determines the support of the estimator (the region within which we combine constraints). It is common to let $g(\vec{x})$ be Gaussian in order to weight constraints in the center of the neighborhood more highly, giving them more influence. The 2D velocity $\hat{u}$ that minimizes $E(\vec{u})$ is the least squares flow estimate".

Result of this estimation is a 2D motion field [108] for all visible surface points which serves as a connection between images of a scene and their relative (pixel-) orientation.

6.1.2 Super-Resolution Construction

Having an estimate about related image orientations (necessarily with sub-pixel accuracy), either by bundle block adjustment or optical flow, the super-resolution can be calculated.

To illustrate the process exemplarily, an 1D-signal is used (see Figure 6.2). There, two segments of the signal are well below the pixel-size (see first row of Figure 6.2 i.e. 70% of the actual sample distance). Rows 2 - 6 show the sampled signal representation when assuming each subsequent position of the line-sensor is shifted by a fifth of a pixel in relation to its predecessor. Subsequently a super-resolved raster needs to be defined. In this example, that raster is five times better resolved than the actual sample distance (see seventh row of Figure 6.2).

The image intensity value at each super-resolved position can be calculated in different ways. One approach is to use the two closest low-resolved intensity values (at opposing sides of the super-pixel of interest) and interpolate linearly (for the 1D-case). For the 2D-case, the approach is extended to a Delaunay-triangulation and the triangle that contains the super-resolved pixel position is used to obtain the intensity values as an intersection of a vector $\vec{v} = (0, 0, 1)^T$ (at the super-resolved pixel position) with the surrounding triangle in 3D-space, where the third dimension is constructed from intensity values of the low-resolved input. This is equivalent to a bi-linear interpolation.

Another technique, which is more resilient to outlier or wrong determined relative image orientation, is inverse distance weighting [109]. There, the image intensity u

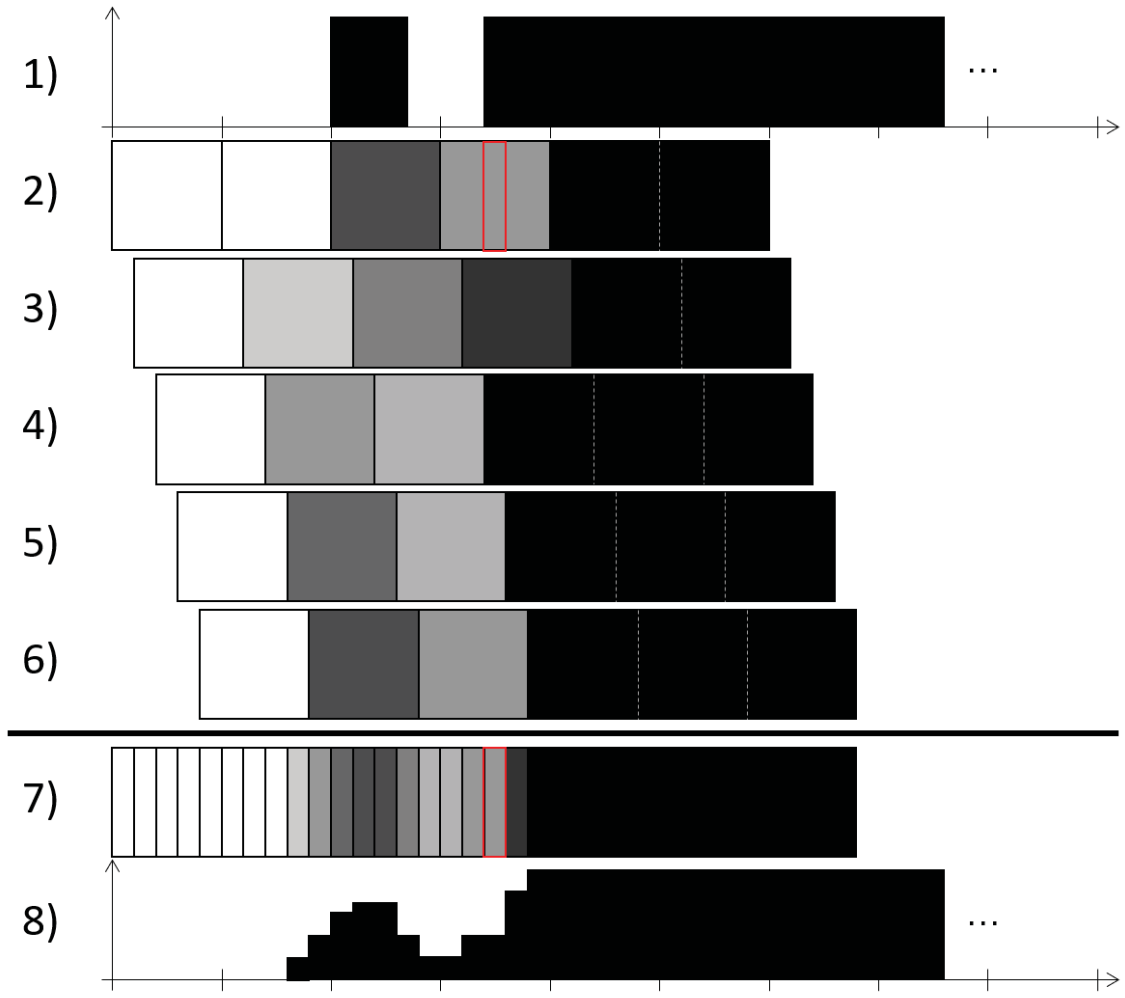


Figure 6.2. Theoretical example of super-resolution, row (1) signal curve with segments below the pixel-size, rows (2) - (6) sampled signal

at super-resolution position x is determined as a weighted sum of all low-resolved intensities u_i divided by the sum of the weights:

$$u(x) = \begin{cases} \frac{\sum_{i=1}^{N} \omega_i(x) \cdot u_i}{\sum_{i=1}^{N} \omega_i(x)}, & \text{if } d(x, x_i) \neq 0 \text{ for all i} \\ u_i, & \text{if } d(x, x_i) = 0 \text{ for some i} \end{cases} \tag{6.4}$$

where every weight is calculated as an reciprocal exponential function according to the euclidean distance between super-resolved pixel position x and the low-resolved position x_i:

$$\omega_i(x) = \frac{1}{d(x, x_i)^p} \tag{6.5}$$

Equation 6.4 allows for a special case, where the super-resolved pixel position x and the low-resolved position x_i is identical and thus the distance $d(x, x_i) = 0$. In that case the super-resolved image intensity value is equal to the low-resolved value u_i.

This special case is of use to illustrate the essence of super-resolution in Figure 6.2. As mentioned before, every line is shifted by exactly a fifth of the original pixel-size starting from accurate aligned position. This implicates that every super-resolved position coincides with one low-resolved position (the two red rectangles in Figure 6.2 show one example) and the super-resolved image intensity values can be constructed from aforementioned special case (row 7).

The final result is given in row 8 and it can be seen that the original signal shape can be reconstructed, although not to the full extend of contrast.

6.2 Demosaicing and Super-Resolution

Impact of the applied demosaicing method has already been outlined in section 4.1. It has been shown that the use of simple conversion methods (e.g. bi-linear interpolation) to transform raw-bayer patterns to color images result in much less spatial resolution than using more sophisticated techniques (e.g. DCB, AHD, see Figure 4.5). This leads to the conclusion that it is greatly beneficial to consider this fact during construction of high-resolved images using multi-view super-resolution.

In 2006, Farsiu et al. [110] published their work regarding multi-view super-resolution and introduced the issue: "In the last two decades, two related categories of problems have been studied independently in image restoration literature: super-resolution and demosaicing. A closer look at these problems reveals the relation between them, and, as conventional color digital cameras suffer from both low-spatial resolution and color-filtering, it is reasonable to address them in a unified context."

Farsiu et al. solve the (low-resolved) image co-registration problem by using optical flow (see section 6.1.1) and assume mostly translational change in the imagery. However, instead of using the L_2-norm during the energy minimization step they propose the use of L_1-norm which proves to be more robust to data outliers [111].

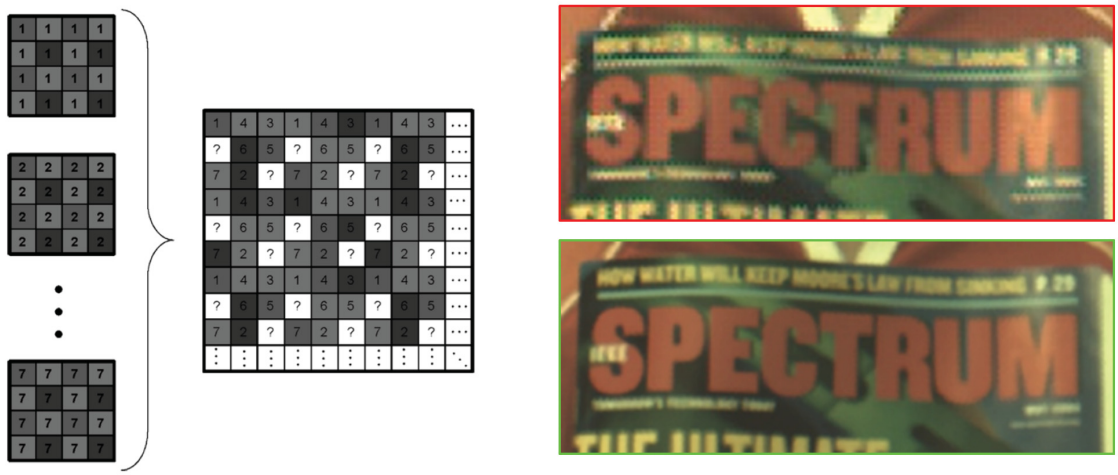


Figure 6.3. "Multiframe Demosaicing and Super-Resolution of Color Images" [110]*, (left) Fusion of seven Bayer pattern LR images with relative translational motion", (top right) one low-resolved image, (bottom right) super-resolved result
*use of this Figure has been approved by and copyright belongs to Prof. Farsiu et al.

Subsequently, a number of relative orientated low-resolved (LR) Bayer-pattern frames is merged into one high-resolved (HR) frame (see Figure 4.5, left). Color matrix of the HR frame (right side of the bracket in Figure 4.5, left) does not follow any Bayer-arrangement and contains empty parts.

Farsiu et al. solve this issue as follows: "We call this operation shift-and-add, which greatly speeds up the task of multiframe image fusion under the assumptions made. To compute the shift-and-add image, first the relative motion between all LR frames is computed. Then, a set of HR images is constructed by up-sampling each LR frame by zero filling. Then, these HR frames are registered with respect to the relative motion of the corresponding LR frames. A pixel-wise mean or median operation on the nonzero values of these HR frames will result in the shift-and-add image."

An exemplary result of this technique is given in Figure 4.5 (top right). There, thirty-three low-resolved images with translational shift have been used to obtain the high-resolved image (see 4.5, bottom right). The improvement becomes obvious when looking at the smallest headline. While it is not readable in the LR example it can be read in the HR variant.

6.3 Digital Surface Models and Super-Resolution

As mentioned before, the use of optical flow as an estimation for relative orientation of low-resolved images to produce super-resolved upgrades is widespread. Furthermore, the basic assumption (or prerequisite) introduced as spatial smoothness of the

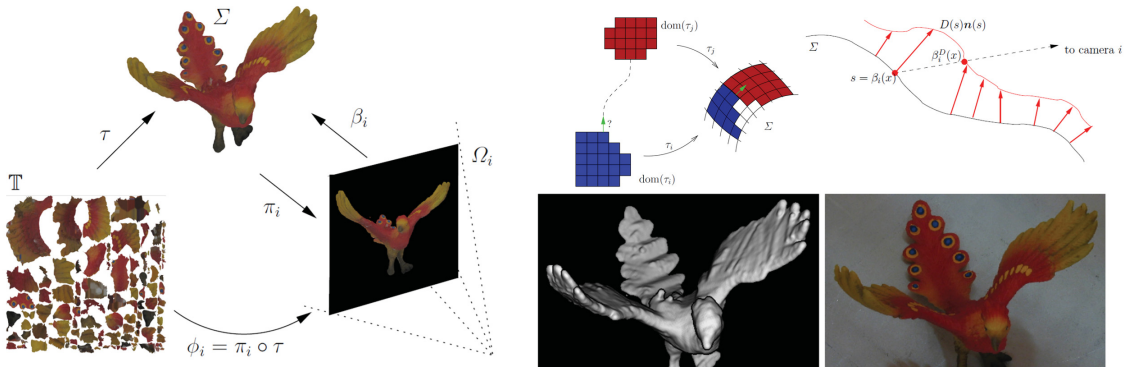


Figure 6.4. "A Super-resolution Framework for High-Accuracy Multiview Reconstruction" [112]*, (left) "The various mappings connecting texture space $\mathbb{T}$, the surface Σ and the image planes Ω_i.", (top right) "Boundary texel neighbour connections [...] on the surface [...] displacement map D", (bottom right) "Intensity-coded backprojection area element $\mathcal{J}_i^{\pi}$ and corresponding input image"
*use of this Figure has been approved by and copyright belongs to Prof. Goldlücke et al.

observed surface often is mandatory [106, 113, 114]. However, this assumption is often not correct and violated especially at object boundaries and edges. Therefore it can be concluded that it would be more precise to take the underlying surface into account when calculating super-resolution.

The work of Goldlücke et al. proposes a solution using pre-reconstructed surfaces during the super-resolution step [112].

They introduce the problem as follows: "In particular, in multi-view settings, usually every patch of the surface is captured from several cameras. Therefore, using a suitable super-resolution model, one should be able to recover the texture map in higher resolution than provided by the input images. However, this possibility has not yet been explored for the curved 3D models obtained by 3D reconstru6ction methods, with most existing methods fitting a local lighting model per-vertex or per-texel only, disregarding texel interdependencies. Thus, our paper aims at opening up the highly interesting area of super-resolution models on surfaces. The super-resolution framework presented in this paper is designed to address all of the shortcomings mentioned above. We account for the interdependency of geometry and photometry by minimizing a single functional with respect to all relevant unknowns: a super-resolved texture map for the surface, a displacement field optimizing local surface geometry, as well as camera calibration parameters."

The single functional, which they set out to minimize, is formulated as:

$$E_{data}(\mathcal{T}, \mathcal{D}, \pi) = \sum_{i=1}^{n} \int_{\Sigma} v_i(\mathcal{J}_i^{\pi} \circ \pi_i)|T - \mathcal{I}_i \circ \pi_i| ds \tag{6.6}$$

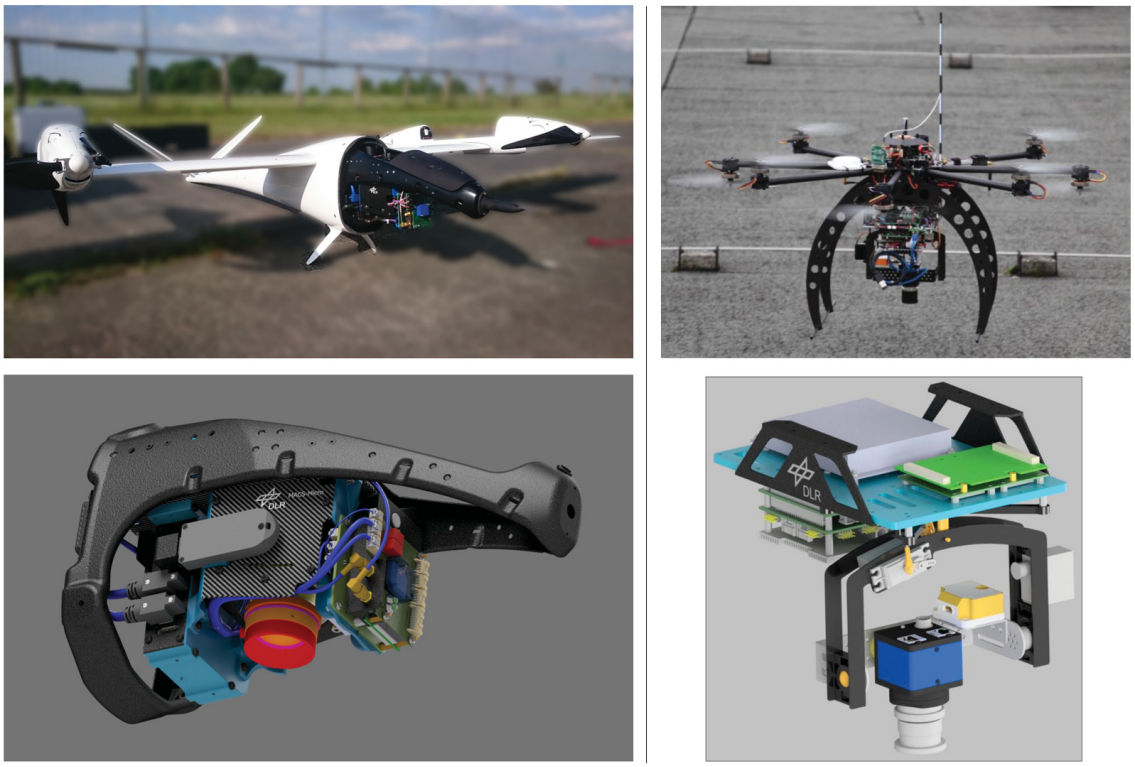


Figure 6.5. Metric camera systems, (top left) Macs Micro 16 MPix with carrier, (bottom left) Macs Micro 16 MPix CAD model, (top right) Macs Micro 12 MPix with carrier, , (bottom right) Macs Micro 12 MPix CAD model

where v_i is a visibility function to determine if a surface point is visible in a source image, $\mathcal{T}$ is the texture space (see Figure 6.4, left), $\mathcal{D}$ the displacement map (see illustration in Figure 6.4, top right) to refine the pre-reconstructed surface, π the projection (camera) parameters (in the context of their work only exterior orientation with 6 DOF but they state that use of interior parameters is straight-forward), Σ the surface, $\mathcal{I}_i$ the input images and $\mathcal{J}_i^{\pi}$ the back-projected images.

In conclusion, the implementation of Goldlücke et al. solve the energy minimization with respect to surface displacement, exterior camera parameters and texture misalignment in one joint functional.

The results given in their publication clearly indicate that it is beneficial to include both the surface and knowledge about the camera to produce super-resolved imagery.

6.4 Structure-Aware Demosaicing and Super-Resolution

Having introduced the essence of super-resolution, as well as two special approaches, simultaneous demosaicing and super-resolution (section 6.2) and super-resolution using surface models and camera-knowledge (section 6.3), a combination of those two techniques is described in this section.

The algorithm employs bundle block adjustment (see section 2.2) to obtain relative orientation parameters instead of optical flow, makes use of precise knowledge about the camera system (see section 6.4.1) and includes derived 3D-surfaces as a result of semi-global matching (see section 2.3.1) and point cloud fusion (see section 2.3.2).

6.4.1 Metric Camera System

Within the German Society of Photogrammetry, Remote Sensing and Geoinformation (DGPF) there was and still is an active debate about UAV based photogrammetric applications. The debate is focused on how such an aerial camera system has to be designed and what key-specifications have to be met [6, 115]. Besides the possibility for direct geo-referencing the main advantage of metric camera systems (and most relevant in super-resolution context) is the higher accuracy of bundle block adjustment under various flight configurations [91, 116].

Therefore, the cameras, later used for structure-aware demosaicing and super-resolution, are briefly introduced and proof of the metric properties is presented exemplary for one camera system. As a side note, both systems took part in the camera resolution benchmark and achieved top-rated results (both MACS Micro variants, see section 5.7, Table 1 and Figure 5.7).

One metric camera system for UAV applications is DLR's MACS Micro 16MP (see Figure 6.5, left). The camera head consists of a 16 MPix CCD sensor (ON Semi-conductor KAI-16070 with Bayer pattern) and an industrial F-Mount lens (Schneider Kreuznach Xenon-Emerald 2.2/50). The aperture is set to f4.0 and the focus is fixed to the hyperfocal distance. Exterior orientation calculation is based on a dual-antenna GNSS receiver (Novatel OEM7720) in combination with an industrial grade MEMS-IMU (Epson G320N). The dual-antenna set-up is used to determine true-heading independently from INS. This improves the orientation accuracy, in particular when movement direction and heading do not correlate due to cross-wind.

Another metric camera system for UAV applications is DLR's MACS Micro 12MP (see Figure 6.5, right). The system consists of an industrial grade camera module with a 12 MPix CMOS chip (4.7μm pixel pitch, Bayer pattern, global shutter mode) and an industrial C-Mount lens (APO Xenoplan f 2.0 24mm). The exterior orientation of each acquired image is obtained by a Dual-Frequency GNSS receiver (Novatel OEM615) combined with an industrial grade IMU (Sensonor STIM 300). The solution is capable of post-process the flight trajectory to gain highest accuracies and every image can be assigned with a GNSS time, GNSS position and attitude. The camera and IMU have been mounted on a stabilized gimbal to compensate the UAV specific flight attitude. The image acquisition, time synchronization, trigger control and co-registration of all data is done by a small embedded PC (Intel Atom Dual Core, 4GB RAM, Linux OS). Due to a continuous syncing of all time counters to the precise GNSS time each image is assigned with a precision less 1μs. The image acquisition is done using a GigE-Vision Interface and the embedded PC has the ability to acquire raw images at a frequency of 4 Hz. All data is recorded in raw

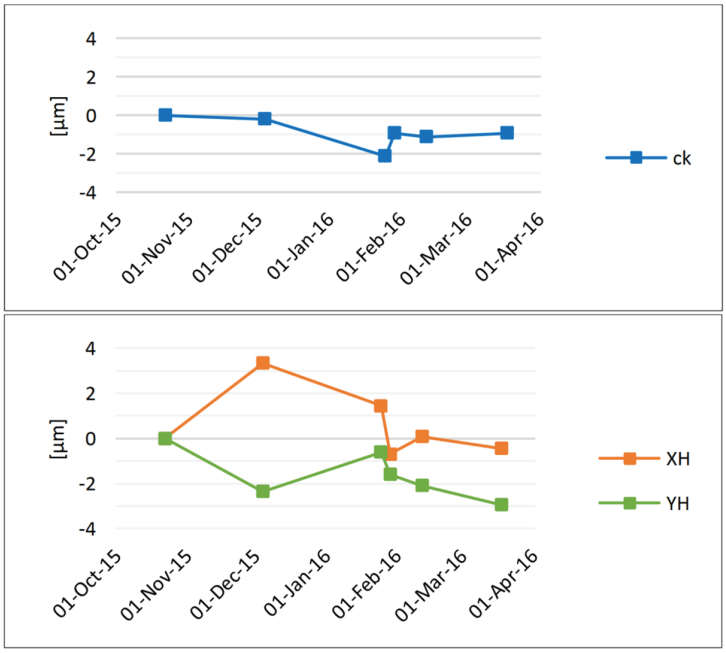


Figure 6.6. Interior orientation parameters for 6 calibrations over a period of 5 months, (top) motion of focal length, (bottom) motion of principal point

(Bayer) format and stored to a swappable Cfast 2.0 Card (up to 256GB).

To prove metric quality of a camera system it needs to be calibrated several times, preferably over a long time span under various weather conditions (especially various temperatures). Under all these conditions, the interior orientation parameters shall be constant [6].

Therefore, the last mentioned system (MACS Micro 12MP) has been calibrated using two different three dimensional test fields located at Beuth University of Applied Sciences in Berlin and Bochum University of Applied Sciences. On both sites, several coded markers are attached to two right-angled facades which serve as ground control points. All markers have been measured with a tachymeter followed by a net adjustment resulting in a standard deviation of 1 mm in each dimension.

The image acquisition was done at different locations, distances, line of sights and height levels in relation to the calibration target. At each location four images with four different rotations were captured (0°, 90°, 180° and 270° around the line of sight axis) considering a homogeneous distribution of markers for each image. The parameters of interior- and exterior orientation have been calculated during a bundle block adjustment using Technet Pictran for the first two calibrations and Aicon 3D-Studio for all following calibrations. The point measurement was done automatically due to the use of coded markers.

The significantly estimated parameters are: calibrated focal length C_K, principal point of auto-collimation X_H and Y_H, radial-symmetric distortion coefficients K_1,

K_2 and misalignment coefficients P_1, P_2.

The calibration procedures have been conducted over a period of time reaching from October to March of the following year at ambient temperatures varying between -3°C and +15°C.

Interior orientation parameters for focal length (C_K) and principal point (X_H and Y_H) plotted over time are given in Figure 6.6.

Due to the fact that the deviations are less than the sensor's pixel pitch (4.7μ m) the interior orientation can be considered as constant. Hence, the chosen industrial grade frame imaging camera clearly shows features of a metric camera and a self-calibration step during flight (in-situ) is not necessary.

Instead, these values can be kept fixed during bundle block adjustment, semi-global matching and during later use as origin of projection for structure-aware demosaicing and super-resolution.

6.4.2 Structure-Aware Demosaicing Algorithm

The final algorithm, described in this section, takes into account and combines numerous of the aforementioned issues and considerations.

It commences with the image acquisition step (see Figure 6.7, a) where images are captured and stored with raw Bayer-pattern information and interior orientation (IO) parameters are well known and long-term stable. Precise exterior orientation (EO) parameters are assigned to every image (geo-coded raw image). These EO parameters are measured and calculated using an industrial grade GNSS/INS receiver including real-time kinematic (RTK) technique, reducing (3-dimensional) EO error to a root mean square error (RMSE) of $\leq$ 3cm [7]. Precise start values for bundle block adjustment (BBA), as a Levenberg-Marquardt minimization (see section 2.2.2), will more likely lead to a solution delivering the global minimum instead of values for one of various local minima, the more accurately start values for the non-linear least square problem are given [117].

In preparation for further processing, the geo-coded raw images are converted to color images (see Figure 6.7, b). However, performances of available demosaicing methods for this color reconstruction process significantly vary (see section 4.1). Therefore, one of the best performing techniques has been chosen for this tool-chain (DCB is set as default).

Subsequently, a bundle block adjustment step is being performed (see Figure 6.7, c) where the IO parameters are confidently kept at fixed pre-calibrated values and simultaneously reducing the number of optimization parameters. The iterative process can be stopped if the overall accuracy (e.g. sigma naught σ_0) is significantly in sub-pixel range ($\sigma_0 \ll 1$ pixel). All results (see section 6.4.3) as outcome of the presented approach are well below 0.5 pixel.

In preparation of dense matching, images are rectified (see Figure 6.7, d) according

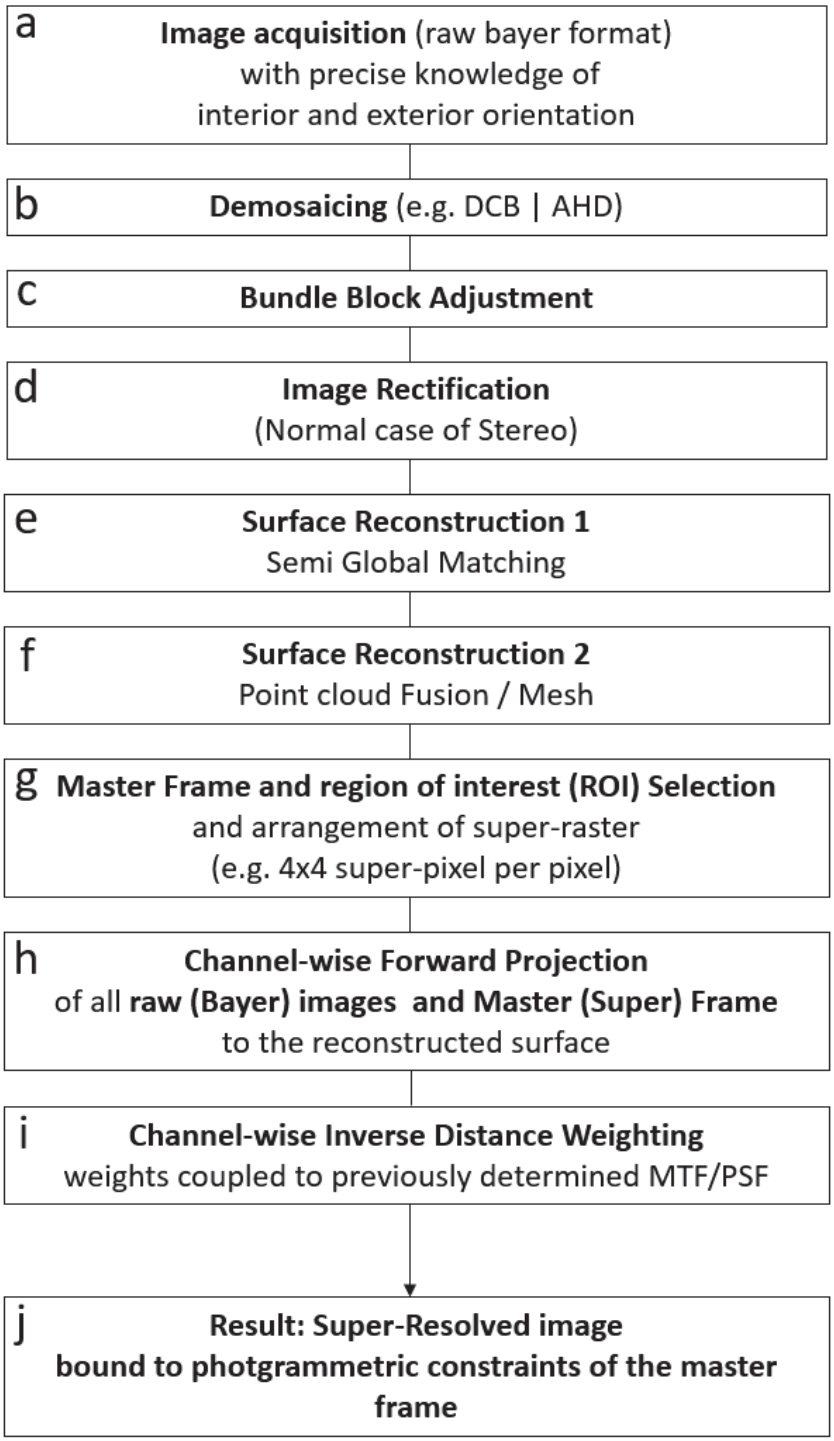


Figure 6.7. Flow diagram of the structure-aware demosaicing algorithm

to the normal case of stereo. "Image rectification is the process of transforming a pair of images [...] in a way that all pairs of homologous iamge points [...] are remapped to the identical row or column in the rectified images [...]." [118]. As a side note, that remapping step includes pixel interpolation and has direct influence on the effective image resolution used for further processing.

Next step is the first part of surface reconstruction (see Figure 6.7, e), the semi global matching [119, 118] (see section 2.3.1). That dense image matching stage delivers disparity maps for every stereo image pair whereby every disparity is converted to depth information and subsequently to 3D world coordinates using precise IO and EO parameters.

Second part of surface reconstruction (see Figure 6.7, f) is the point cloud fusion and subsequent mesh respectively digital surface model (DSM) generation (see section 2.3.2).

Next (intermediate) step of structure-aware demosaicing is to select a region of interest (ROI) for which the resolution should be improved (see Figure 6.7, g). The current implemented state of the presented algorithm takes a 3D-shape (e.g. polygon) as input. Furthermore, the selection of one or more master-frames needs to be done in this step. These master-frames are the images where the super-resolution will be modelled into. There are several criteria to select a master frame. One criterion is the nadir-angle. There, the image having and angle (between a ray from camera to ROI-center and the surface) closest to 90 degree is selected. If a resolution target is depicted in the images, the best resolved image could be selected as master frame. Another criterion, when looking at iterative reconstruction improvement (see section 6.4.4) is the base-height ratio [120]. Then, images fulfilling the best parameter constellation are selected as master frames. Completing this step, every pattern of (one or more) master frame(s) is divided into a super-raster (e.g. 4×4 super-pixel per pixel, see Figure 6.8).

Precise 3D-reconstruction of the observed scene features the next step of structure-aware demosaicing (see Figure 6.7, h). There, the super-pixel footprint of the previously selected master frame is (forward) projected onto the reconstructed surface. In contrast to back-projection of world coordinates to image plane (see equation 2.24), the forward projection is represented as a ray (λ) from camera- to world coordinates when transposing aforementioned equation to [86]:

$$X(\lambda) = P^{+}x + \lambda C \tag{6.7}$$

where P^{+} is the pseudo-inverse of projection matrix P with $P^{+} = P^{T}(PP^{T})^{-1}$. The final 3D world coordinate is calculated as the ray's intersection with the surface (e.g. ray-tracing). The current state-of implementation, presented here, divides the surface into piecewise planar representations and performs a test for every ray and surface patch if both intersect inside the patch's polygon. Subsequently, the found (forward-projected) world coordinate of every (master-frame) super-pixel is

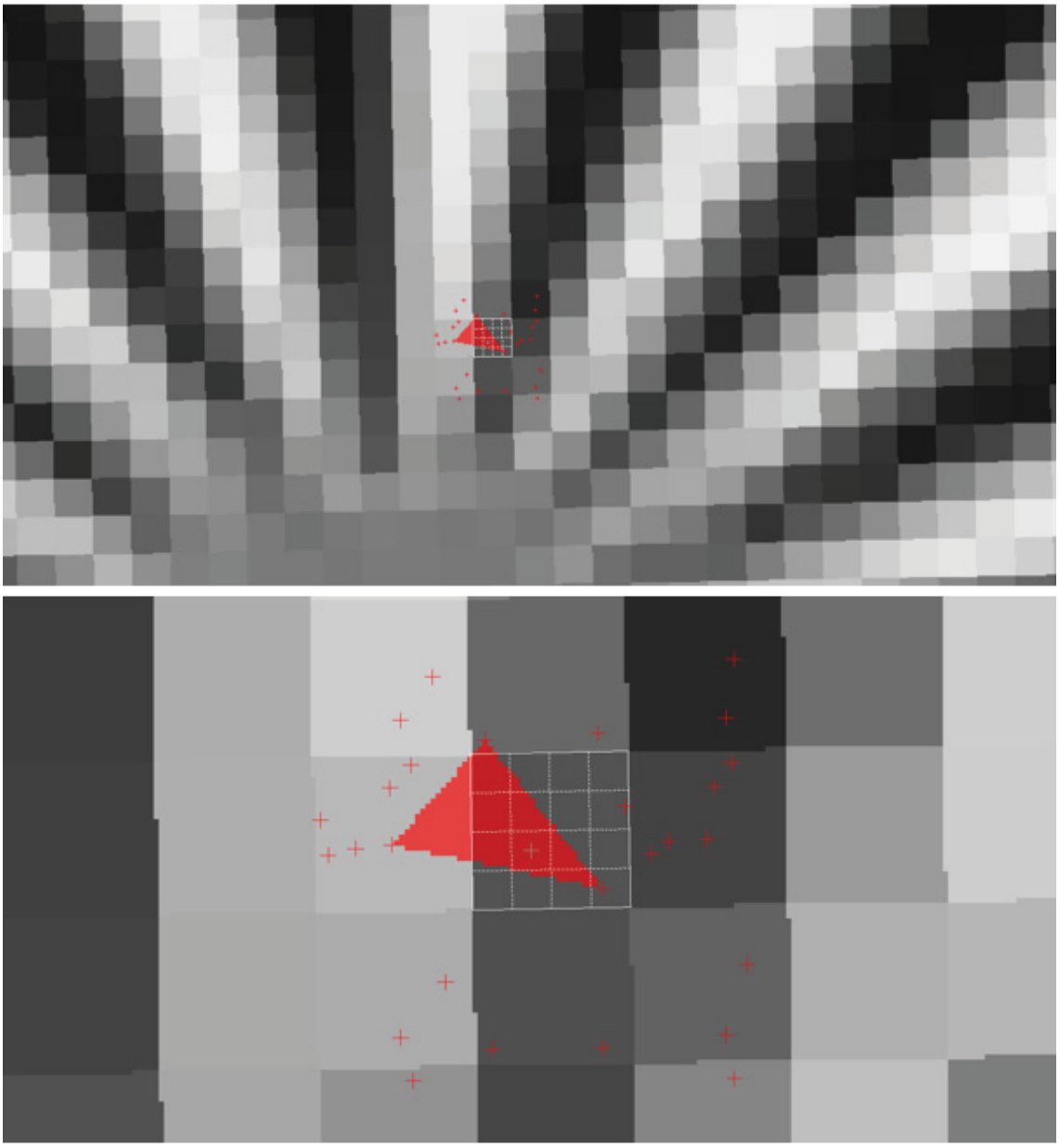


Figure 6.8. (top background) master frame projection onto the surface (top foreground) projection of all relevant red pixels of all contributing images, grey squares show projection of the master-frame including super-raster; red triangle constructed from the three closest, enclosing pixels; grey cross is the current super-pixel position (bottom) zoomed depiction of the above mentioned

back-projected into every image designated for the super-resolution process. Also, the distorted image coordinates are calculated applying equation 2.26 and 2.27. These back-projected and distorted coordinates will most likely be not identical to a pixel center. However, the final channel-wise forward projection uses the pixel center of the found pixels and all centers of its neighbourhood (e.g 3×3 neighbourhood) located in the raw-images (Bayer-pattern). Before calculating the forward-projection ray (equation 6.7) the identified pixels are grouped according to their channel (red, green, blue) and the distortion is being reversed. The reversal of equations 2.26 and 2.27 can not be solved directly but as an iterative process [121]. Result of this step of the presented super-resolution method is a channel-wise projection of all raw-images (including the master frame) onto the reconstructed surface (see Figure 6.8).

Final step is the super-resolution calculation itself (see Figure 6.7, i). There, for every channel at every super-pixel position the corresponding intensity value is calculated according to the inverse distance weighting approach (equation 6.4). Exploiting knowledge of the imaging-system's (previously determined) achievable resolution the maximum allowed distance $d_{max}(x, x_i)$ is determined as follows:

$$d_{max}(x, x_i) = r_{MTF10} + \sigma_0 \tag{6.8}$$

where the distance r_{MTF10} is the reciprocal of k_{MTF10} (spatial frequency to MTF at 10% modulation, see section 2.1.2) and σ_0 the overall accuracy measure of previously performed BBA. In other words, only pixel having a distance smaller than or equal to $d_{max}(x, x_i)$ are included in the super-resolution calculation. As mentioned before, instead of inverse distance weighting, it would also be possible to apply a Delaunay-triangulation including subsequent bi-linear interpolation (see Figure 6.8, red triangle) to obtain super resolved images. However, all results presented in the following section (6.4.3) show that inverse distance weighting slightly but consistently performs better compared to Delaunay-triangulation (~3% better).

6.4.3 Results

The method, presented and developed as part of this thesis, to determine spatial resolution objectively, can now be applied to quantify resolution improvement. Therefore, values for k_{MTF10} and r_{MTF10} (see section 2.1.2) are calculated for input images as well as for the super-resolved image and compared subsequently. Additionally, if a bar test-target is imaged (e.g. USAF51 target) both resolution values (input and super-resolution) can be confirmed visually (see section 3.1.1).

The first result, presented here, is obtained using images of the MACS Micro 16MP camera system with CCD sensor (see section 6.4.1 and 5.7). Eight low-resolved aerial images have been utilized to calculate one super-resolved image (see Figure 6.9) and values for spatial resolution have been determined for both cases. Values for k_{MTF10} (unit is line/pixel) of the best low-resolved input image (used demosaicing method DCB) are 0.888, 0.868, 0.883 for red, green and blue channel. Reciprocal of these values multiplied by ground sample distance (input image GSD is 1.24cm) delivers

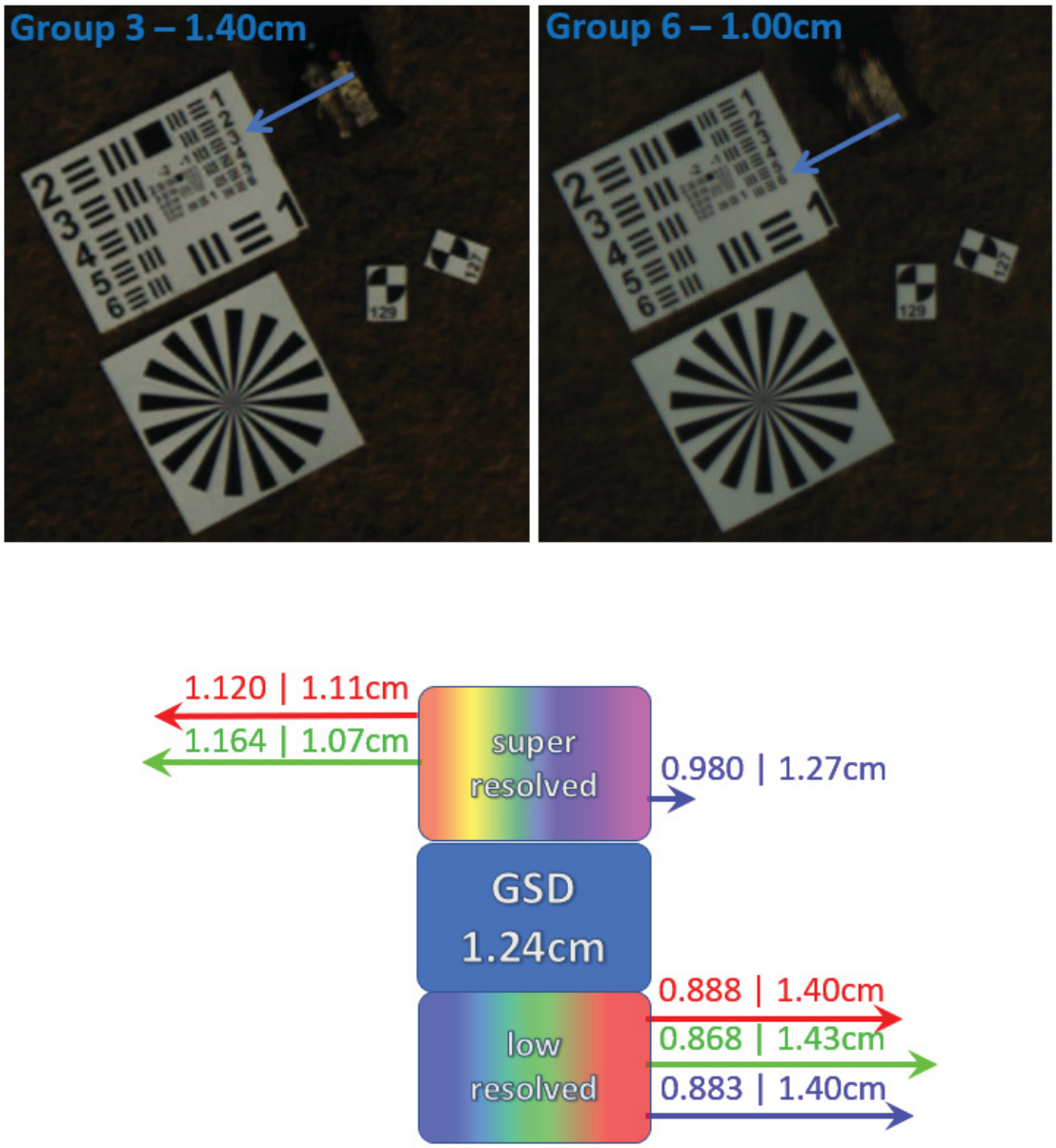


Figure 6.9. Result for structure-aware demosaicing, Camera system MACS Micro 16MP CCD (upper left) best resolved input image, least discriminable group identified by 10 operators group 3 (upper right) super resolved image, least discriminable group identified by 10 operators group 6 (bottom) resolution comparison of low- and super-resolved images in relation to GSD

effective image resolution r_{MTF10} and may be confirmed using the bar test-target. In that case the effective spatial resolution is 1.39cm, 1.42cm and 1.39cm (red, green, blue). The measurement has been approved by 10 operators. They identified group 3 as the least discriminable cluster of the displayed bar-test target (see Figure 6.9, upper left) which is equivalent to 1.40cm.

Super-resolution has been calculated using inverse distance weighting (distance weight exponential $p = 12$, see equation 6.5). Measured spatial resolution is 1.120, 1.164 and 0.980 for red, green and blue channel. Again, reciprocal of these values multiplied by ground sample distance (1.24cm) delivers effective image resolution, which then is calculated to 1.11cm, 1.07cm and 1.27cm (red, green, blue). The operator driven resolution determination (10 operators) delivered group 6 (see Figure 6.9, upper right) as the least discriminable cluster of the displayed (super-resolved) bar-test target which is equivalent to 1.00cm.

Averaging resolution values of all three channels of the super-resolved image gives 1.15cm which is well below calculated ground sample distance but slightly worse than human observers would assess.

When comparing resolution between best low-resolved (1.40cm) and super-resolved (1.15cm) image directly, the resolution improvement is approximately 18% (see Figure 6.9, bottom).

The second result, presented here, is obtained using images of the MACS Micro 12MP camera system with CMOS sensor (see section 6.4.1 and 5.7). Six low-resolved aerial images have been utilized to calculate one super-resolved image (see Figure 6.10) and values for spatial resolution have been determined for both cases. Values for k_{MTF10} (unit is line/pixel) of the best low-resolved input image (used demosaicing method DCB) are 0.898, 0.882, 0.895 for red, green and blue channel. Ground sample distance for this flight and the best low-resolved image is 1.05cm. Once again, reciprocal of k_{MTF10} multiplied by GSD delivers effective image resolution r_{MTF10} and may be confirmed using the bar test-target. In that case the effective spatial resolution is 1.17cm, 1.19cm and 1.17cm (red, green, blue). The measurement has been approved by 10 operators. They identified group 4 as the least discriminable cluster of the displayed bar-test target (see Figure 6.10, upper left) which is equivalent to 1.25cm.

Super-resolution has been calculated using inverse distance weighting (distance weight exponential $p = 12$, see equation 6.5). Measured spatial resolution is 0.981, 1.092 and 0.959 for red, green and blue channel. Again, reciprocal of these values multiplied by ground sample distance (1.05cm) delivers effective image resolution, which in then is calculated to 1.07cm, 0.96cm and 1.09cm (red, green, blue). The operator driven resolution determination (10 operators) delivered group 6 (see Figure 6.10, upper right) as the least discriminable cluster of the displayed (super-resolved) bar-test target which is equivalent to 1.00cm.

Averaging resolution values of all three channels of the super-resolved image gives 1.04cm which is just about equal to calculated ground sample distance and also equal

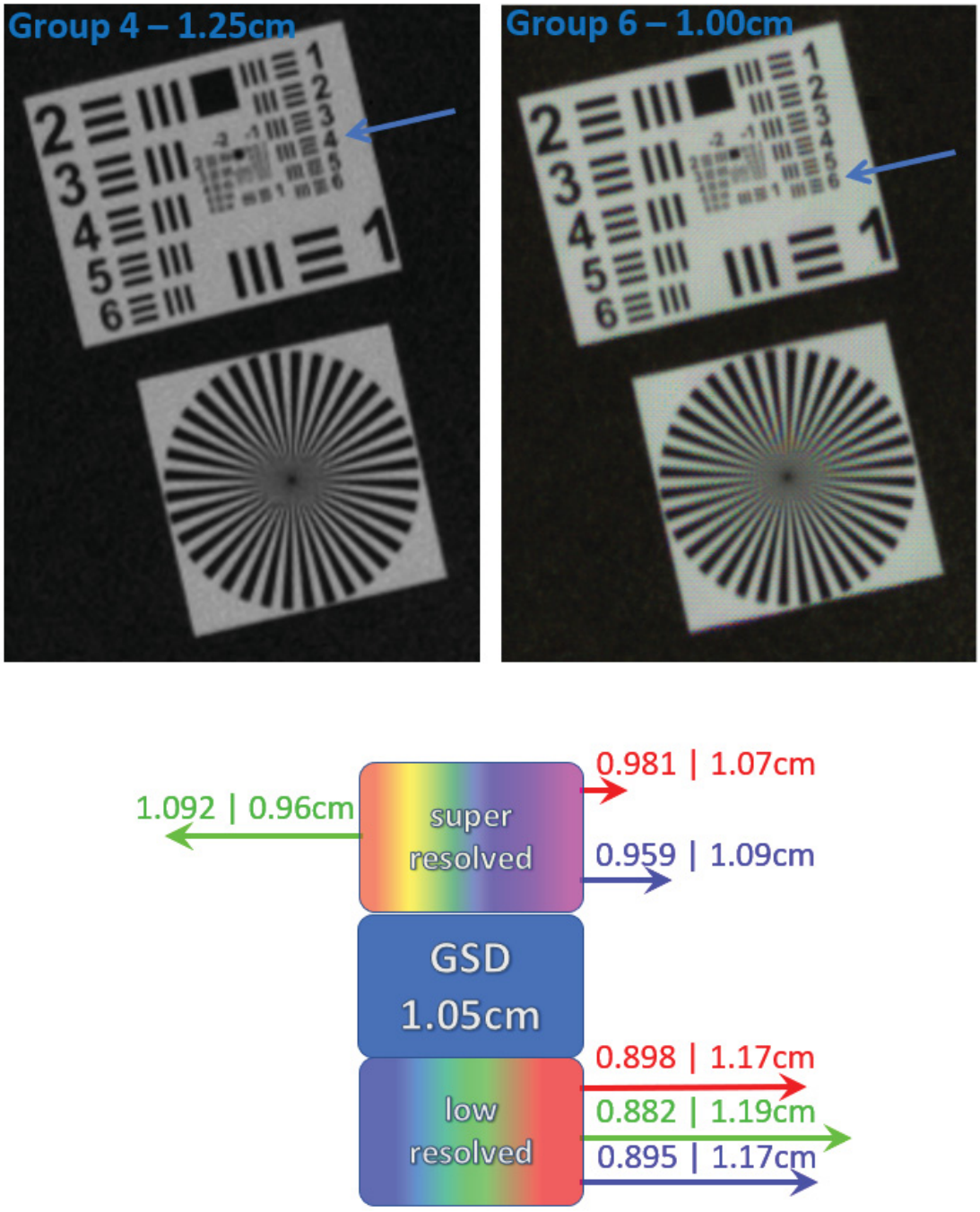


Figure 6.10. Result for structure-aware demosaicing, Camera system MACS Micro 12MP CMOS (upper left) best resolved input image, least discriminable group identified by 10 operators group 4 (upper right) super resolved image, least discriminable group identified by 10 operators group 6 (bottom) resolution comparison of low- and super-resolved images in relation to GSD

to the assessment of human observers.

When comparing resolution between best low-resolved (1.18cm) and super-resolved (1.04cm) image directly, the resolution improvement is approximately 12% (see Figure 6.10, bottom).

Two observations shall be emphasized at this point. First, resolving power can be improved applying the presented super resolution technique, even beyond the calculated ground sample distance. Although, resolution improvement between 12%-18% still leaves room for further enhancements (e.g. $\geq 25\%$). The second observation concerns improvement of the different color channels. In both examples, blue channel was the least enhanced channel. Assumed reason can be that both focal length and radial-symmetric distortion parameter are calibrated for one mean wave-length. An extended camera model (e.g. one for each channel) may cause more significant enhancements and is part of future work and investigations (see following section 6.4.4).

The number of images, used for super resolution, plays an important role. The results presented by Farsiu et al. [110] (see Figure 6.3) for example included 31 low-resolved images. Aerial image campaigns with in-track image overlap between 80% and 90% provide 5-10 overlapping images suitable for super-resolution calculation. Utilizing images from adjacent flight stripes will not enhance the results due to higher in-stripe block-stability than inter-stripe block-stability. That statement is strongly supported by the amount of valid matches per stereo image pair. When looking at the disparity maps (result of semi global matching) of images from one stripe compared to matching results of images from adjacent flight-stripes, the amount of valid matchings in-stripe is much higher than the amount of valid matchings inter-stripe. This circumstance in turn, can be traced back to BBA where in-stripe features are more frequent and simultaneously determined more robustly than inter-stripe features.

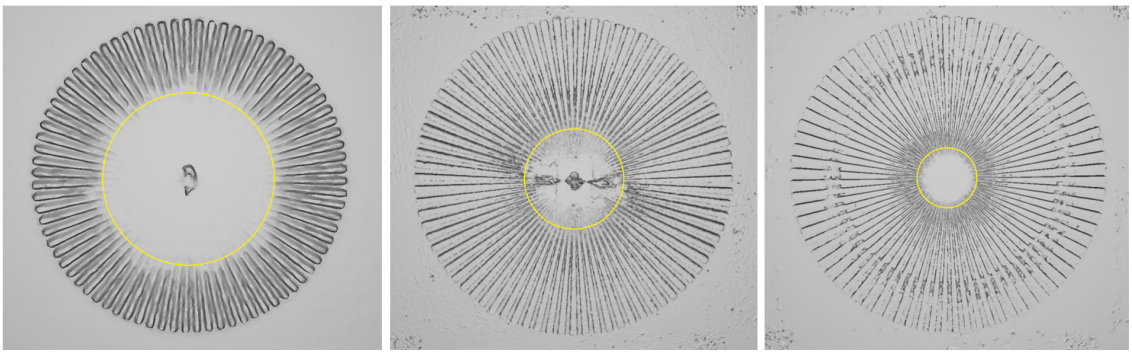


Figure 6.11. 3D reconstruction of a Böhler-star (left) using low-resolved simulated images (middle) using two enlarged and bi-cubic interpolated images (right) using two super-resolved images

6.4.4 Future Work

As mentioned in previous section, extending the camera model, especially different focal length and different radial-symmetric distortion parameters for every channel is one point that will be investigated as part of future work. Obtained results strongly indicate that this would further improve the outcome.

Another aspect, that possibly can be applied on top of the camera model changes, is to merge all parameters into one energy minimization step, very similar to the approach proposed by Goldlücke et al. [112] but with an extended camera model and complete 3D-reconstruction instead of surface displacement estimation.

The hypothesis, that super-resolved images, when maintaining photogrammetric constrains, may even produce better results for 3D-surface reconstruction, can be supported as part of a simulation. Therefore, images of a synthetic 3D scene including a Böhler-star [122] have been rendered using pre-defined (highly accurate) EO, IO and world-scene parameters. The Böhler-star is an extension of the Siemens-star to third dimension (depth) and the simulater has been designed to validate photogrammetric algorithms [85].

Applying the whole 3D-surface reconstruction work flow (see section 2.3) delivers a fused and meshed point cloud having a perceptible spatial (now three-dimensional) resolution limit (see Figure 6.11, left). Also, 3D-reconstruction of super-resolved images using simulated input has been obtained and the (3D) resolution is noticeably better compared to low-resolved input (see Figure 6.11, right). However, both semi global matching and point cloud fusion implementations use pre-defined filter-size parameters, often coupled to GSD. Instead of comparing low-resolved 3D-reconstruction to super-resolved 3D-reconstruction, the comparison is made between super-resolved 3D-reconstruction and the 3D-surface derived from images transformed to equal pixel size as the super-resolved reference frame and simultaneously applying bi-cubic interpolation (see Figure 6.11, middle). As a result, filter-size parameters are equal

and still the super-resolved reconstruction is noticeably better compared to bi-cubic enlarged low-resolved input.

It strongly indicates that the potential to combine all parameters (EO, IO, 3D-representation) into one iterative energy minimization step should enhance both, spatial image resolution and world scene (3D) resolution, even though highly accurate EO, IO and world-scene parameters have been used as super-resolution input. Augmenting the implementation with that technique and feasibility study on real data is also part of future work.

7 Summary and Conclusion

The present thesis provided a description of the imaging process of remote sensing sensors mathematically and methodologically including imaging components and subsequent processes with the aim of obtaining spatial resolution parameters objectively. It has been shown that repeatable and reliable resolution determination only can be achieved by understanding and including all parts of the imaging process, starting with the sensor components and external circumstances through to all involved pre-processing and post-processing steps.

Several techniques to determine resolution parameters have been presented and reviewed out of which an approach using a Siemens-star (spokes) target has been proven to be best fit. The identified standardization aspects which need to be considered are summarized as follows: used demosaicing method, applied signal scan interpolation method, MTF approximation technique, Siemens-star center position, normalization of contrast magnitude, influence of test target inclination and motion blur.

Main outcome of all findings is a software tool that will be official part of the upcoming norm DIN 18740-8 "Photogrammetric products – Part 8: Requirements for image quality (quality of optical remote sensing data)" released by the German Institute for Standardization.

Reliability of the software tool has been verified by several self-validating procedures, namely measurements for different exposure times, measurements for different number of Siemens-star segments, measurements for sinusoidal and square wave Siemens-stars, comparison to other techniques (Slanted-edge, bar test target), measurements of predefined model-based MTF parameters. Furthermore, repeatability and independence of subjective influence has been proven.

Several benchmark procedures have been presented to support the decision making process finding best suited sensor-lens combinations for different designated applications. A camera resolution benchmark has been described and conducted for ten different sensor-lens combinations under laboratory conditions. A similar procedure has been described for operating conditions and has been compared to laboratory results. Furthermore, an application delivering a complete image-field MTF has been developed to provide a total resolution representation of the considered camera.

Final chapter of the thesis on hand presented a super-resolution approach to improve spatial resolution of aerial images. The essence and state of the art of multi-view super-resolution has been described followed by two specific super-resolution approaches.

A further software tool has been implemented that is able to combine aforementioned two different super-resolution techniques. The tool includes known image quality parameters, raw imagery (non-demosaiced) and derived digital surface models in subsequent calculations and the magnitude of resolution improvement has been quantified with the aforementioned software tool for standardized measurements.

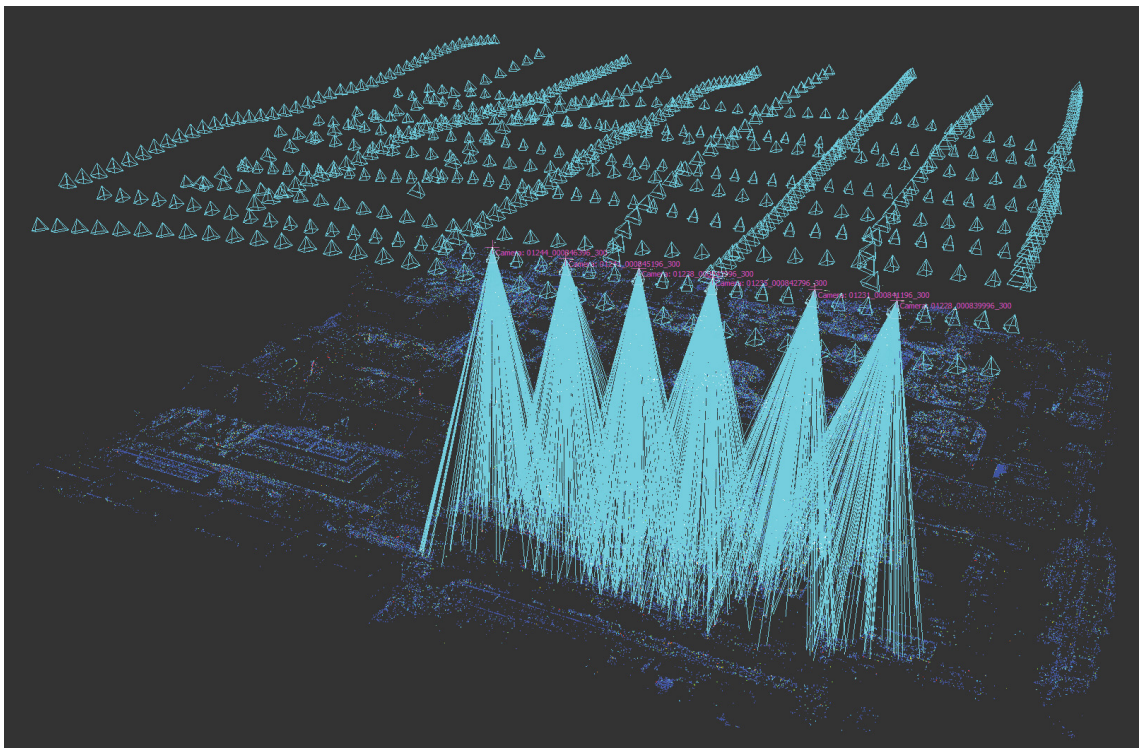


Figure 8.1. Block configuration

8 Appendix

8.1 Correlation of Extracted Features and Bundle Block Adjustment

As stated in section 6.4.1 the camera is primary sensor for further photogrammetric processing as it delivers the observations of which the later products (tie points for bundle adjustment, sparse and dense 3D point clouds) are being derived. To support this, a real UAV-based data set from DLRs MACS Micro 16MP camera system (see section 6.4.1 and 5.7) has been acquired over a photogrammetric test site.

8.1.1 Simulation Setup

A bundle block adjustment was calculated for all 632 images in cross-flight configuration (overlap: 80% in flight direction, 60% across flight direction, see Figure 8.1).

A high precise geodetic network with 45 signalized points has been applied out of which 5 points served as control points and the remaining 40 points were used as check points. In a second run the original 632 images have been blurred through Gaussian blur filter (see Figure 8.2 upper left and upper right) and the bundle block was recalculated.

Nr. of features per Image (mean)	39569	Difference %	Nr. of features per Image (mean)	36315
		8.2		
Matched features per Image (mean)	21706	13.7	Matched features per Image (mean)	18739

RMSE [cm]	Check points		RMSE [cm]	Check points
X-Error	0.349		X-Error	0.337
Y-Error	0.236		Y-Error	0.247
Z-Error	1.372	19.9	Z-Error	1.712
3D-Error	1.435	18.6	3D-Error	1.762
projection error (mean)[pix]	0.408		projection error (mean)[pix]	0.415

Figure 8.2. Correlation of extracted features and bundle block adjustment, (top left) visual impression of original image, (top right) the same image blurred by Gaussian filter, (bottom) Simulated effects of image sharpness on bundle block adjustment for original images, blurred images and corresponding differences

8.1.2 Results

It can be seen that both, the number of detected features and check point accuracy, decrease. Figure 8.2 (bottom) compares both results.

While total amount of automatic extracted SIFT-features decreases by 8.2% the amount of valid features used for bundle block adjustment diminishes by 13.7% from original to blurred images. Resulting 3D-Error differs by 18.6% with Z-Error component (19.9%) as main factor of influence. Based on the assumption that real lens systems not only are degraded through Gaussian blur but also aberration effects, the real differences could assumingly be even higher than within this simulation.

8.2 Protocols

8.2.1 MTF Measurement in Relation to Applied Demosaicing Method

Table 2. MTF measurement in relation to applied demosaicing method differences are given in Table 3

ImageID	channel	AHD method	DCB method	VNG method	MHC method	BiLin method
MM121638	red	1.002	0.993	0.926	0.896	0.678
	green	0.956	0.948	0.854	0.869	0.827
	blue	0.992	0.996	0.927	0.917	0.601
MM121755	red	0.927	0.927	0.893	0.856	0.628
	green	0.901	0.886	0.834	0.839	0.799
	blue	0.927	0.928	0.900	0.891	0.572
MM121951	red	0.966	0.967	0.942	0.890	0.632
	green	0.936	0.923	0.888	0.882	0.850
	blue	0.970	0.967	0.944	0.911	0.632
MM163212	red	0.963	0.956	0.910	0.928	0.476
	green	0.923	0.904	0.835	0.849	0.837
	blue	0.946	0.943	0.895	0.854	0.653
MM165618	red	1.217	1.239	1.112	1.146	0.659
	green	1.146	1.156	1.014	1.068	0.994
	blue	1.174	1.207	1.095	1.132	0.714
MM165712	red	1.113	1.141	1.046	1.046	0.721
	green	1.068	1.073	0.891	0.941	0.824
	blue	1.124	1.134	1.009	1.000	0.665
MM163515	red	0.994	1.003	0.948	0.928	0.662
	green	0.949	0.954	0.864	0.867	0.781
	blue	0.985	0.996	0.947	0.924	0.680
MM163621	red	1.114	1.096	1.010	1.010	0.669
	green	1.000	0.998	0.948	0.955	0.898
	blue	1.015	1.070	1.011	0.989	0.736
MM163717	red	0.819	0.830	0.792	0.785	0.491
	green	0.793	0.791	0.748	0.759	0.744
	blue	0.812	0.816	0.790	0.754	0.565
MM163848	red	0.470	0.469	0.422	0.449	0.375
	green	0.462	0.457	0.445	0.460	0.448
	blue	0.450	0.457	0.433	0.448	0.388
MM165754	red	1.064	1.117	1.004	0.986	0.666
	green	0.994	0.991	0.940	0.946	0.871
	blue	1.008	1.029	0.973	0.964	0.678
MM165834	red	1.194	1.225	1.126	1.151	0.620
	green	1.154	1.175	1.065	1.067	0.993
	blue	1.193	1.212	1.108	1.112	0.691
IP176	red	0.692	0.675	0.666		0.444
	green	0.681	0.662	0.653		0.639
	blue	0.654	0.649	0.639		0.518
IP187	red	0.655	0.647	0.629		0.519
	green	0.655	0.637	0.619		0.595
	blue	0.630	0.632	0.564		0.450
IP198	red	0.667	0.674	0.638		0.558
	green	0.654	0.632	0.619		0.608
	blue	0.637	0.624	0.602		0.452
P1436	red	0.685	0.645	0.631		0.493
	green	0.694	0.630	0.642		0.642
	blue	0.713	0.670	0.668		0.418
P1447	red	0.621	0.600	0.561		0.436
	green	0.603	0.585	0.555		0.561
	blue	0.667	0.617	0.584		0.401
P1452	red	0.582	0.575	0.537		0.449
	green	0.576	0.568	0.551		0.561
	blue	0.567	0.566	0.549		0.369
P1472	red	0.476	0.478	0.441		0.387
	green	0.496	0.506	0.482		0.491
	blue	0.494	0.507	0.437		0.343
P1492	red	0.636	0.635	0.609		0.439
	green	0.617	0.608	0.586		0.577
	blue	0.678	0.673	0.647		0.447

ImageID	channel	AHD method	DCB method	VNG method	MHC method	BiLin method
PM217	red	0.821	0.756	0.727		0.606
	green	0.792	0.678	0.684		0.633
	blue	0.772	0.681	0.668		0.460
PM226	red	0.676	0.622	0.608		0.386
	green	0.695	0.684	0.630		0.660
	blue	0.658	0.670	0.616		0.480
PM235	red	0.766	0.710	0.714		0.547
	green	0.741	0.705	0.696		0.674
	blue	0.724	0.716	0.698		0.495
R7671	red	0.633	0.633	0.608		0.523
	green	0.611	0.596	0.582		0.553
	blue	0.646	0.644	0.612		0.548
R7672	red	0.842	0.794	0.803		0.639
	green	0.818	0.724	0.740		0.669
	blue	0.822	0.774	0.778		0.607
R7673	red	1.115	1.056	1.004		0.695
	green	1.045	0.944	0.910		0.830
SA265	red	0.395	0.402	0.368		0.354
	green	0.377	0.364	0.384		0.372
	blue	0.370	0.377	0.363		0.339

Table 3. Deviation to VNG approach in percent with values given in Table 2

ImageID	channel	Δ (AHD - VNG) [%]	Δ (DCB - VNG) [%]	Δ (MHC - VNG) [%]	Δ (LIN - VNG) [%]
MM121638	red	7.6	6.7	-3.3	-26.8
	green	10.7	9.9	1.7	-3.2
	blue	6.6	6.9	-1.1	-35.2
MM121755	red	3.7	3.7	-4.3	-29.7
	green	7.4	5.9	0.6	-4.2
	blue	2.9	3.0	-1.0	-36.4
MM121951	red	2.5	2.6	-5.8	-32.9
	green	5.1	3.8	-0.7	-4.3
	blue	2.7	2.4	-3.6	-33.1
MM163212	red	5.5	4.8	1.9	-47.7
	green	9.5	7.6	1.6	0.2
	blue	5.4	5.1	-4.8	-27.0
MM165618	red	8.6	10.3	3.0	-40.7
	green	11.5	12.3	5.1	-2.0
	blue	6.7	9.3	3.3	-34.8
MM165712	red	6.0	8.3	0.0	-31.1
	green	16.6	17.0	5.3	-7.5
	blue	10.2	11.0	-0.9	-34.1
MM163515	red	4.6	5.5	-2.2	-30.2
	green	9.0	9.4	0.3	-9.6
	blue	3.9	4.9	-2.5	-28.2
MM163621	red	9.3	7.8	0.0	-33.8
	green	5.2	5.0	0.7	-5.3
	blue	0.4	5.5	-2.2	-27.2
MM163717	red	3.3	4.6	-0.9	-38.0
	green	5.7	5.4	1.4	-0.5
	blue	2.7	3.2	-4.8	-28.5
MM163848	red	10.2	10.0	6.0	-11.1
	green	3.7	2.6	3.3	0.7
	blue	3.8	5.3	3.3	-10.4
MM165754	red	5.6	10.1	-1.8	-33.7
	green	5.4	5.1	0.6	-7.3
	blue	3.5	5.4	-0.9	-30.3
MM165834	red	5.7	8.1	2.2	-44.9
	green	7.7	9.4	0.2	-6.8
	blue	7.1	8.6	0.4	-37.6
IP176	red	3.8	1.3		-33.3
	green	4.1	1.4		-2.1
	blue	2.3	1.5		-18.9
IP187	red	4.0	2.8		-17.5
	green	5.5	2.8		-3.9
	blue	10.5	10.8		-20.2
IP198	red	4.3	5.3		-12.5
	green	5.4	2.1		-1.8
	blue	5.5	3.5		-24.9
P1436	red	7.9	2.2		-21.9
	green	7.5	-1.9		0.0
	blue	6.3	0.3		-37.4
P1447	red	9.7	6.5		-22.3
	green	8.0	5.1		1.1
	blue	12.4	5.3		-31.3
P1452	red	7.7	6.6		-16.4
	green	4.3	3.0		1.8
	blue	3.2	3.0		-32.8
P1472	red	7.4	7.7		-12.2
	green	2.8	4.7		1.9
	blue	11.5	13.8		-21.5
P1492	red	4.2	4.1		-27.9
	green	5.0	3.6		-1.5
	blue	4.6	3.9		-30.9
PM217	red	11.4	3.8		-16.6
	green	13.6	-0.9		-7.5
	blue	13.5	1.9		-31.1

ImageID	channel	Δ (AHD - VNG) [%]	Δ (DCB - VNG) [%]	Δ (MHC - VNG) [%]	Δ (LIN - VNG) [%]
PM226	red	10.1	2.3		-36.5
	green	9.4	7.9		4.8
	blue	6.4	8.1		-22.1
PM235	red	6.8	-0.6		-23.4
	green	6.1	1.3		-3.2
	blue	3.6	2.5		-29.1
R7671	red	3.9	3.9		-14.0
	green	4.7	2.3		-5.0
	blue	5.3	5.0		-10.5
R7672	red	4.6	-1.1		-20.4
	green	9.5	-2.2		-9.6
	blue	5.4	-0.5		-22.0
R7673A	red	10.0	4.9		-30.8
	green	12.9	3.6		-8.8
	blue	6.3	0.0		-30.2
SA265	red	6.8	8.5		-3.8
	green	-1.9	-5.5		-3.1
	blue	1.9	3.7		-6.6
Overall	red	**6.5**	**5.2**	**-0.4**	**-26.3**
	green	**7.2**	**4.5**	**1.7**	**-3.2**
	blue	**5.7**	**4.9**	**-1.2**	**-27.1**

8.2.2 MTF Measurement in Relation to Applied Scan Method

Table 4. MTF measurement in relation to applied signal scan method

ImageID	channel	near. neighbor (NN)	bi-linear (BL)	bi-cubic (BC)	Δ(BL-NN) [%]	Δ(BL-BC) [%]
MM161708D	red	0.922	0.827	0.899	10.3	8.0
	green	0.880	0.810	0.865	8.0	6.4
	blue	0.918	0.828	0.900	9.8	8.0
MM161709D	red	0.944	0.857	0.918	9.2	6.6
	green	0.928	0.840	0.897	9.5	6.4
	blue	0.947	0.866	0.925	8.6	6.4
MM161710D	red	0.984	0.924	0.968	6.1	4.5
	green	0.969	0.903	0.951	6.8	5.0
	blue	0.992	0.925	0.972	6.8	4.8
MM161711D	red	0.897	0.817	0.864	8.9	5.4
	green	0.889	0.813	0.857	8.5	5.1
	blue	0.896	0.825	0.873	7.9	5.5
MM161712D	red	0.819	0.755	0.798	7.8	5.4
	green	0.818	0.761	0.799	7.0	4.8
	blue	0.831	0.773	0.809	7.0	4.4
MM161713D	red	0.763	0.718	0.749	5.9	4.1
	green	0.758	0.719	0.749	5.1	4.0
	blue	0.763	0.725	0.759	5.0	4.5
MM161714D	red	0.817	0.773	0.810	5.4	4.6
	green	0.821	0.783	0.816	4.6	4.0
	blue	0.830	0.796	0.827	4.1	3.7
MM161715D	red	0.930	0.868	0.913	6.7	4.9
	green	0.894	0.843	0.887	5.7	5.0
	blue	0.890	0.839	0.883	5.7	5.0
MM161716D	red	0.953	0.886	0.939	7.0	5.6
	green	0.930	0.867	0.925	6.8	6.3
	blue	0.947	0.884	0.939	6.7	5.9
MM161717D	red	0.929	0.848	0.910	8.7	6.8
	green	0.913	0.841	0.896	7.9	6.1
	blue	0.931	0.857	0.915	7.9	6.3
RC7673D16	red	1.118	1.010	1.064	9.7	5.1
	green	0.991	0.916	0.985	7.6	7.0
	blue	1.007	0.955	1.005	5.2	5.0
RC7673D36	red	1.085	0.990	1.028	8.8	3.7
	green	1.008	0.883	0.985	12.4	10.4
	blue	1.025	0.965	1.006	5.9	4.1
RC7697D	red	0.831	0.793	0.824	4.6	3.8
	green	0.791	0.762	0.797	3.7	4.4
	blue	0.832	0.795	0.823	4.4	3.4
RC7699V	red	0.940	0.846	0.914	10.0	7.4
	green	0.888	0.815	0.869	8.2	6.2
	blue	0.930	0.845	0.911	9.1	7.2
MM163212M	red	0.997	0.922	0.968	7.5	4.8
	green	0.923	0.848	0.909	8.1	6.7
	blue	0.908	0.856	0.903	5.7	5.2
MM163212D16	red	0.461	0.463	0.479	0.4	3.3
	green	0.865	0.831	0.878	3.9	5.4
	blue	0.688	0.650	0.668	5.5	2.7
MM163212D36	red	0.523	0.515	0.528	1.5	2.5
	green	0.851	0.796	0.843	6.5	5.6
	blue	0.681	0.657	0.685	3.5	4.1
MM163212D16	red	1.043	0.960	0.996	8.0	3.6
	green	0.974	0.908	0.954	6.8	4.8
	blue	1.005	0.946	0.983	5.9	3.8
MM163212D36	red	1.025	0.976	1.010	4.8	3.4
	green	0.973	0.913	0.950	6.2	3.9
	blue	1.012	0.959	0.995	5.2	3.6
PM218D	red	1.009	0.864	0.941	14.4	8.2
	green	0.962	0.838	0.909	12.9	7.8
	blue	0.881	0.811	0.865	7.9	6.2

ImageID	channel	near. neighbor (NN)	bi-linear (BL)	bi-cubic (BC)	Δ(BL-NN) [%]	Δ(BL-BC) [%]
PM219D	red	0.939	0.870	0.932	7.3	6.7
	green	0.892	0.839	0.903	5.9	7.1
	blue	0.848	0.812	0.869	4.2	6.6
PM220D	red	0.859	0.793	0.844	7.7	6.0
	green	0.794	0.736	0.795	7.3	7.4
	blue	0.755	0.715	0.762	5.3	6.2
PO5984D	red	0.919	0.861	0.922	6.3	6.6
	green	0.872	0.826	0.876	5.3	5.7
	blue	0.967	0.880	0.934	9.0	5.8
PO5984L	red	0.654	0.621	0.647	5.0	4.0
	green	0.819	0.779	0.828	4.9	5.9
	blue	0.552	0.530	0.550	4.0	3.6
PO5983V	red	0.898	0.814	0.878	9.4	7.3
	green	0.829	0.769	0.833	7.2	7.7
	blue	0.919	0.848	0.916	7.7	7.4
PM183D	red	0.820	0.783	0.824	4.5	5.0
	green	0.791	0.744	0.794	5.9	6.3
	blue	0.795	0.738	0.793	7.2	6.9
PM208D	red	0.669	0.661	0.686	1.2	3.6
	green	0.631	0.630	0.654	0.2	3.7
	blue	0.629	0.626	0.653	0.5	4.1
MM121951L	red	0.678	0.637	0.665	6.0	4.2
	green	0.924	0.846	0.895	8.4	5.5
	blue	0.654	0.633	0.658	3.2	3.8
MM12370D	red	0.823	0.778	0.823	5.5	5.5
	green	0.817	0.771	0.816	5.6	5.5
	blue	0.835	0.783	0.829	6.2	5.5
MM12373D	red	1.022	0.935	0.992	8.5	5.7
	green	0.991	0.917	0.953	7.5	3.8
	blue	1.039	0.937	1.005	9.8	6.8
MM12367D	red	1.003	0.937	0.992	6.6	5.5
	green	0.965	0.886	0.954	8.2	7.1
	blue	0.996	0.927	0.986	6.9	6.0
SA9259D	red	1.049	0.945	0.982	9.9	3.8
	green	0.916	0.815	0.866	11.0	5.9
	blue	0.992	0.914	0.955	7.9	4.3
SIM050	red	1.242	1.225	1.243	1.4	1.4
	green	1.242	1.225	1.243	1.4	1.4
	blue	1.242	1.225	1.243	1.4	1.4
SIM075	red	0.830	0.798	0.835	3.9	4.4
	green	0.830	0.798	0.835	3.9	4.4
	blue	0.830	0.798	0.835	3.9	4.4
SIM100	red	0.644	0.618	0.644	4.0	4.0
	green	0.644	0.618	0.644	4.0	4.0
	blue	0.644	0.618	0.644	4.0	4.0
SIM125	red	0.505	0.507	0.530	0.4	4.3
	green	0.505	0.507	0.530	0.4	4.3
	blue	0.505	0.507	0.530	0.4	4.3
SIM150	red	0.445	0.443	0.450	0.4	1.6
	green	0.445	0.443	0.450	0.4	1.6
	blue	0.445	0.443	0.450	0.4	1.6
SIM175	red	0.385	0.380	0.385	1.3	1.3
	green	0.385	0.380	0.385	1.3	1.3
	blue	0.385	0.380	0.385	1.3	1.3
SIMO10	red	0.649	0.620	0.648	4.5	4.3
	green	0.649	0.620	0.648	4.5	4.3
	blue	0.649	0.620	0.648	4.5	4.3
MM16431D	red	0.819	0.834	0.855	1.8	2.5
	green	0.811	0.830	0.845	2.3	1.8
	blue	0.829	0.843	0.869	1.7	3.0
MM16432D	red	0.909	0.831	0.879	8.6	5.5
	green	0.887	0.821	0.859	7.4	4.4
	blue	0.930	0.835	0.888	10.2	6.0
MM16433D	red	0.948	0.847	0.914	10.7	7.3
	green	0.920	0.838	0.896	8.9	6.5
	blue	0.967	0.867	0.932	10.3	7.0

ImageID	channel	near. neighbor (NN)	bi-linear (BL)	bi-cubic (BC)	Δ(BL-NN) [%]	Δ(BL-BC) [%]
MM16434D	red	0.999	0.904	0.961	9.5	5.9
	green	0.994	0.881	0.944	11.4	6.7
	blue	1.008	0.900	0.959	10.7	6.2
MM16435D	red	1.065	0.958	1.011	10.0	5.2
	green	1.010	0.922	0.977	8.7	5.6
	blue	1.061	0.962	1.012	9.3	4.9
MM16436D	red	0.951	0.898	0.946	5.6	5.1
	green	0.914	0.867	0.919	5.1	5.7
	blue	0.952	0.898	0.945	5.7	5.0
MM16437D	red	0.866	0.802	0.845	7.4	5.1
	green	0.861	0.796	0.838	7.5	5.0
	blue	0.880	0.814	0.860	7.5	5.3
MM16553D	red	0.710	0.681	0.709	4.1	3.9
	green	0.703	0.670	0.701	4.7	4.4
	blue	0.701	0.674	0.704	3.9	4.3
MM16554D	red	0.700	0.666	0.706	4.9	5.7
	green	0.693	0.658	0.699	5.1	5.9
	blue	0.707	0.666	0.707	5.8	5.8
MM16555D	red	0.539	0.508	0.532	5.8	4.5
	green	0.504	0.488	0.504	3.2	3.2
	blue	0.509	0.495	0.512	2.8	3.3
MM16556D	red	0.527	0.509	0.527	3.4	3.4
	green	0.502	0.485	0.501	3.4	3.2
	blue	0.509	0.488	0.504	4.1	3.2
Overall	**red**				**6.2**	**4.8**
	green				**6.1**	**5.2**
	blue				**5.8**	**4.8**

8.2.3 MTF Measurement in Relation to Applied MTF Approximation

Table 5. MTF Measurement in relation to applied MTF approximation method

ImageID	channel	piecewise linear (L)	polynomial (P)	gaussian (G)	Δ(L-P) [%]	Δ(L-G) [%]
PM176D	red	0.674	0.633	0.549	6.1	18.5
	green	0.662	0.620	0.567	6.3	14.4
	blue	0.648	0.577	0.484	11.0	25.3
PM183D	red	0.780	0.791	0.609	1.4	21.9
	green	0.743	0.724	0.559	2.6	24.8
	blue	0.735	0.635	0.411	13.6	44.1
PM187D	red	0.647	0.610	0.542	5.7	16.2
	green	0.636	0.604	0.545	5.0	14.3
	blue	0.630	0.584	0.495	7.3	21.4
PM196D	red	0.591	0.544	0.350	8.0	40.8
	green	0.604	0.568	0.408	6.0	32.5
	blue	0.618	0.601	0.384	2.8	37.9
PM198D	red	0.672	0.681	0.640	1.3	4.8
	green	0.635	0.630	0.570	0.8	10.2
	blue	0.622	0.586	0.478	5.8	23.2
PM208D	red	0.666	0.669	0.546	0.5	18.0
	green	0.634	0.634	0.530	0.0	16.4
	blue	0.630	0.595	0.429	5.6	31.9
PM243D	red	0.683	0.674	0.589	1.3	13.8
	green	0.661	0.652	0.611	1.4	7.6
	blue	0.695	0.689	0.669	0.9	3.7
PM254D	red	0.556	0.512	0.498	7.9	10.4
	green	0.591	0.561	0.535	5.1	9.5
	blue	0.611	0.583	0.559	4.6	8.5
PM265D	red	0.705	0.657	0.647	6.8	8.2
	green	0.693	0.670	0.642	3.3	7.4
	blue	0.696	0.685	0.677	1.6	2.7
SA257D	red	0.492	0.438	0.330	11.0	32.9
	green	0.444	0.426	0.307	4.1	30.9
	blue	0.425	0.423	0.239	0.5	43.8
SA258D	red	0.717	0.705	0.674	1.7	6.0
	green	0.645	0.596	0.549	7.6	14.9
	blue	0.695	0.683	0.494	1.7	28.9
SA259D	red	0.939	0.907	1.084	3.4	15.4
	green	0.810	0.803	0.751	0.9	7.3
	blue	0.909	0.876	0.999	3.6	9.9
CA003D	red	0.897	0.931	0.662	3.8	26.2
	green	0.838	0.821	0.696	2.0	16.9
	blue	0.855	0.861	0.747	0.7	12.6
CA004D	red	0.870	0.935	0.616	7.5	29.2
	green	0.826	0.827	0.693	0.1	16.1
	blue	0.872	0.861	0.723	1.3	17.1
CA010D	red	0.432	0.427	0.289	1.2	33.1
	green	0.661	0.657	0.797	0.6	20.6
	blue	0.582	0.302	0.309	48.1	46.6
RC671D	red	0.610	0.586	0.507	3.9	16.9
	green	0.561	0.530	0.454	5.5	19.1
	blue	0.599	0.563	0.466	6.0	22.2
RC672D	red	0.789	0.791	0.768	0.3	2.7
	green	0.720	0.738	0.644	2.5	10.6
	blue	0.766	0.782	0.689	2.1	10.1
RC673D	red	1.005	0.906	1.008	9.9	0.3
	green	0.912	0.885	0.760	3.0	16.7
	blue	0.948	0.892	0.925	5.9	2.4
PO982D	red	0.746	0.762	0.577	2.1	22.7
	green	0.734	0.718	0.565	2.2	23.0
	blue	0.761	0.791	0.433	3.9	43.1
PO983D	red	0.797	0.822	0.675	3.1	15.3
	green	0.733	0.787	0.659	7.4	10.1
	blue	0.833	0.916	0.558	10.0	33.0

ImageID	channel	piecewise linear (L)	polynomial (P)	gaussian (G)	Δ(L-P) [%]	Δ(L-G) [%]
PO984D	red	0.862	0.861	0.746	0.1	13.5
	green	0.832	0.827	0.702	0.6	15.6
	blue	0.886	0.945	0.604	6.7	31.8
PO985D	red	0.915	0.909	0.809	0.7	11.6
	green	0.839	0.842	0.728	0.4	13.2
	blue	0.921	0.934	0.718	1.4	22.0
PO986D	red	0.697	0.684	0.480	1.9	31.1
	green	0.688	0.656	0.509	4.7	26.0
	blue	0.707	0.687	0.463	2.8	34.5
PM176L	red	0.444	0.451	0.450	1.6	1.4
	green	0.647	0.577	0.543	10.8	16.1
	blue	0.519	0.459	0.426	11.6	17.9
PM183L	red	0.540	0.540	0.487	0.0	9.8
	green	0.709	0.650	0.532	8.3	25.0
	blue	0.466	0.420	0.370	9.9	20.6
PM187L	red	0.521	0.480	0.449	7.9	13.8
	green	0.597	0.561	0.521	6.0	12.7
	blue	0.448	0.446	0.419	0.4	6.5
PM196L	red	0.406	0.387	0.302	4.7	25.6
	green	0.579	0.545	0.389	5.9	32.8
	blue	0.518	0.465	0.328	10.2	36.7
PM198L	red	0.551	0.532	0.489	3.4	11.3
	green	0.609	0.590	0.544	3.1	10.7
	blue	0.445	0.438	0.399	1.6	10.3
PM208L	red	0.541	0.496	0.439	8.3	18.9
	green	0.615	0.594	0.501	3.4	18.5
	blue	0.442	0.424	0.377	4.1	14.7
PM243L	red	0.514	0.486	0.465	5.4	9.5
	green	0.621	0.602	0.570	3.1	8.2
	blue	0.547	0.536	0.531	2.0	2.9
PM254L	red	0.460	0.457	0.444	0.7	3.5
	green	0.554	0.523	0.504	5.6	9.0
	blue	0.500	0.471	0.455	5.8	9.0
PM265L	red	0.528	0.495	0.516	6.3	2.3
	green	0.673	0.619	0.605	8.0	10.1
	blue	0.495	0.516	0.544	4.2	9.9
SA257L	red	0.354	0.358	0.295	1.1	16.7
	green	0.471	0.445	0.316	5.5	32.9
	blue	0.318	0.307	0.225	3.5	29.2
SA258L	red	0.525	0.483	0.454	8.0	13.5
	green	0.658	0.614	0.567	6.7	13.8
	blue	0.453	0.455	0.400	0.4	11.7
SA259L	red	0.687	0.602	0.580	12.4	15.6
	green	0.891	0.823	0.795	7.6	10.8
	blue	0.700	0.650	0.633	7.1	9.6
CA003L	red	0.438	0.439	0.419	0.2	4.3
	green	0.896	0.832	0.783	7.1	12.6
	blue	0.547	0.539	0.493	1.5	9.9
CA004L	red	0.486	0.439	0.416	9.7	14.4
	green	0.916	0.884	0.785	3.5	14.3
	blue	0.518	0.531	0.504	2.5	2.7
CA010L	red	0.279	0.298	0.263	6.8	5.7
	green	0.734	0.711	0.863	3.1	17.6
	blue	0.223	0.246	0.243	10.3	9.0
RC671L	red	0.509	0.467	0.393	8.3	22.8
	green	0.535	0.505	0.431	5.6	19.4
	blue	0.452	0.474	0.391	4.9	13.5
RC672L	red	0.636	0.599	0.538	5.8	15.4
	green	0.659	0.678	0.596	2.9	9.6
	blue	0.590	0.558	0.506	5.4	14.2
RC673L	red	0.686	0.675	0.623	1.6	9.2
	green	0.822	0.823	0.686	0.1	16.5
	blue	0.675	0.663	0.619	1.8	8.3
PO982L	red	0.539	0.505	0.420	6.3	22.1
	green	0.660	0.632	0.509	4.2	22.9
	blue	0.459	0.444	0.351	3.3	23.5

ImageID	channel	piecewise linear (L)	polynomial (P)	gaussian (G)	Δ(L-P) [%]	Δ(L-G) [%]
PO983L	red	0.592	0.571	0.504	3.5	14.9
	green	0.735	0.727	0.600	1.1	18.4
	blue	0.472	0.469	0.415	0.6	12.1
PO984L	red	0.611	0.589	0.530	3.6	13.3
	green	0.795	0.774	0.623	2.6	21.6
	blue	0.535	0.495	0.438	7.5	18.1
PO985L	red	0.655	0.622	0.551	5.0	15.9
	green	0.779	0.776	0.650	0.4	16.6
	blue	0.590	0.555	0.475	5.9	19.5
PO986L	red	0.532	0.505	0.407	5.1	23.5
	green	0.669	0.634	0.491	5.2	26.6
	blue	0.516	0.506	0.395	1.9	23.4
PM176V	red	0.664	0.525	0.482	20.9	27.4
	green	0.653	0.608	0.596	6.9	8.7
	blue	0.637	0.529	0.452	17.0	29.0
PM183V	red	0.781	0.679	0.533	13.1	31.8
	green	0.728	0.688	0.583	5.5	19.9
	blue	0.739	0.525	0.387	29.0	47.6
PM187V	red	0.620	0.553	0.500	10.8	19.4
	green	0.614	0.596	0.580	2.9	5.5
	blue	0.565	0.505	0.447	10.6	20.9
PM196V	red	0.415	0.421	0.304	1.4	26.7
	green	0.545	0.528	0.395	3.1	27.5
	blue	0.543	0.502	0.335	7.6	38.3
PM198V	red	0.637	0.620	0.561	2.7	11.9
	green	0.619	0.623	0.597	0.6	3.6
	blue	0.601	0.522	0.444	13.1	26.1
PM208V	red	0.622	0.573	0.487	7.9	21.7
	green	0.622	0.618	0.554	0.6	10.9
	blue	0.603	0.493	0.400	18.2	33.7
PM243V	red	0.675	0.603	0.524	10.7	22.4
	green	0.647	0.645	0.634	0.3	2.0
	blue	0.661	0.614	0.584	7.1	11.6
PM254V	red	0.567	0.502	0.489	11.5	13.8
	green	0.565	0.543	0.542	3.9	4.1
	blue	0.553	0.506	0.484	8.5	12.5
PM265V	red	0.705	0.612	0.576	13.2	18.3
	green	0.682	0.650	0.660	4.7	3.2
	blue	0.676	0.587	0.569	13.2	15.8
SA257V	red	0.436	0.388	0.311	11.0	28.7
	green	0.443	0.438	0.325	1.1	26.6
	blue	0.374	0.341	0.226	8.8	39.6
SA258V	red	0.756	0.669	0.524	11.5	30.7
	green	0.655	0.632	0.612	3.5	6.6
	blue	0.701	0.564	0.430	19.5	38.7
SA259V	red	0.966	1.064	0.726	10.1	24.8
	green	0.838	0.805	0.868	3.9	3.6
	blue	0.935	0.948	0.769	1.4	17.8
CA003V	red	0.894	1.083	0.450	21.1	49.7
	green	0.781	0.782	0.862	0.1	10.4
	blue	0.854	0.952	0.573	11.5	32.9
CA004V	red	0.864	1.094	0.433	26.6	49.9
	green	0.789	0.789	0.855	0.0	8.4
	blue	0.860	1.045	0.543	21.5	36.9
R7673A	red	1.115	1.096	1.018	1.7	8.7
	green	1.045	1.036	0.862	0.9	17.5
	blue	1.040	1.040	0.879	0.0	15.5
RC671V	red	0.557	0.527	0.430	5.4	22.8
	green	0.560	0.548	0.469	2.1	16.3
	blue	0.571	0.544	0.423	4.7	25.9
RC672V	red	0.800	0.792	0.625	1.0	21.9
	green	0.730	0.733	0.668	0.4	8.5
	blue	0.763	0.717	0.570	6.0	25.3
RC673V	red	1.009	1.043	0.797	3.4	21.0
	green	0.906	0.857	0.789	5.4	12.9
	blue	0.975	1.033	0.747	5.9	23.4

ImageID	channel	piecewise linear (L)	polynomial (P)	gaussian (G)	Δ(L-P) [%]	Δ(L-G) [%]
PO982L	red	0.724	0.714	0.516	1.4	28.7
	green	0.712	0.711	0.584	0.1	18.0
	blue	0.780	0.883	0.392	13.2	49.7
PO983L	red	0.809	0.819	0.557	1.2	31.1
	green	0.773	0.769	0.656	0.5	15.1
	blue	0.840	1.044	0.460	24.3	45.2
PO984L	red	0.869	0.923	0.602	6.2	30.7
	green	0.819	0.794	0.690	3.1	15.8
	blue	0.903	1.066	0.491	18.1	45.6
PO985L	red	0.889	1.010	0.650	13.6	26.9
	green	0.846	0.818	0.745	3.3	11.9
	blue	0.939	1.071	0.587	14.1	37.5
PO986L	red	0.705	0.623	0.460	11.6	34.8
	green	0.678	0.662	0.536	2.4	20.9
	blue	0.725	0.683	0.439	5.8	39.4
MM161638D	red	0.987	0.967	1.010	2.0	2.3
	green	0.944	0.925	0.920	2.0	2.5
	blue	0.986	0.986	0.950	0.0	3.7
MM161755D	red	0.936	0.950	0.935	1.5	0.1
	green	0.887	0.905	0.843	2.0	5.0
	blue	0.932	0.976	0.836	4.7	10.3
MM161951D	red	0.964	0.950	1.070	1.5	11.0
	green	0.923	0.911	0.974	1.3	5.5
	blue	0.962	0.957	1.036	0.5	7.7
MM163212D	red	0.967	1.018	0.956	5.3	1.1
	green	0.915	0.917	0.908	0.2	0.8
	blue	0.951	0.958	0.964	0.7	1.4
MM163515D	red	0.996	1.017	0.897	2.1	9.9
	green	0.947	0.962	0.820	1.6	13.4
	blue	0.986	1.025	0.794	4.0	19.5
MM163621D	red	1.007	1.063	1.163	5.6	15.5
	green	0.969	0.971	1.028	0.2	6.1
	blue	1.004	1.016	1.132	1.2	12.7
MM161638L	red	0.669	0.698	0.677	4.3	1.2
	green	0.827	0.822	0.786	0.6	5.0
	blue	0.595	0.589	0.592	1.0	0.5
MM161755L	red	0.642	0.626	0.614	2.5	4.4
	green	0.810	0.805	0.733	0.6	9.5
	blue	0.579	0.563	0.540	2.8	6.7
MM161951L	red	0.630	0.636	0.653	1.0	3.7
	green	0.847	0.845	0.840	0.2	0.8
	blue	0.630	0.669	0.665	6.2	5.6
MM163212L	red	0.472	0.514	0.538	8.9	14.0
	green	0.837	0.843	0.800	0.7	4.4
	blue	0.653	0.698	0.691	6.9	5.8
MM163515L	red	0.653	0.632	0.569	3.2	12.9
	green	0.771	0.795	0.695	3.1	9.9
	blue	0.667	0.644	0.533	3.4	20.1
MM163621L	red	0.680	0.673	0.671	1.0	1.3
	green	0.903	0.844	0.852	6.5	5.6
	blue	0.743	0.725	0.722	2.4	2.8
MM161638V	red	0.934	0.948	0.862	1.5	7.7
	green	0.864	0.842	0.882	2.5	2.1
	blue	0.930	1.034	0.739	11.2	20.5
MM161755V	red	0.885	0.912	0.776	3.1	12.3
	green	0.835	0.824	0.808	1.3	3.2
	blue	0.892	0.993	0.690	11.3	22.6
MM161951V	red	0.935	0.973	0.843	4.1	9.8
	green	0.885	0.861	0.939	2.7	6.1
	blue	0.937	0.971	0.881	3.6	6.0
MM163212V	red	0.916	1.091	0.669	19.1	27.
	green	0.838	0.836	0.881	0.2	5.1
	blue	0.902	0.946	0.864	4.9	4.2
MM163515V	red	0.940	0.994	0.703	5.7	25.2
	green	0.854	0.833	0.771	2.5	9.7
	blue	0.935	1.003	0.641	7.3	31.4

ImageID	channel	piecewise linear (L)	polynomial (P)	gaussian (G)	Δ(L-P) [%]	Δ(L-G) [%]
MM163621V	red	1.005	1.079	0.894	7.4	11.0
	green	0.945	0.881	0.941	6.8	0.4
	blue	1.004	1.038	0.912	3.4	9.2
MM161708D	red	0.822	0.827	0.756	0.6	8.0
	green	0.807	0.812	0.711	0.6	11.9
	blue	0.823	0.840	0.711	2.1	13.6
MM161709D	red	0.858	0.861	0.860	0.3	0.2
	green	0.841	0.841	0.822	0.0	2.3
	blue	0.868	0.888	0.793	2.3	8.6
MM161710D	red	0.923	0.911	0.887	1.3	3.9
	green	0.901	0.884	0.889	1.9	1.3
	blue	0.924	0.948	0.807	2.6	12.7
MM161711D	red	0.815	0.816	0.788	0.1	3.3
	green	0.812	0.808	0.835	0.5	2.8
	blue	0.824	0.843	0.727	2.3	11.8
MM161712D	red	0.780	0.781	0.720	0.1	7.7
	green	0.781	0.783	0.780	0.3	0.1
	blue	0.794	0.823	0.672	3.7	15.4
MM161713D	red	0.724	0.726	0.693	0.3	4.3
	green	0.724	0.741	0.756	2.3	4.4
	blue	0.734	0.761	0.643	3.7	12.4
MM161714D	red	0.794	0.791	0.724	0.4	8.8
	green	0.795	0.792	0.782	0.4	1.6
	blue	0.803	0.833	0.666	3.7	17.1
MM161715D	red	0.879	0.876	0.812	0.3	7.6
	green	0.849	0.839	0.844	1.2	0.6
	blue	0.846	0.892	0.723	5.4	14.5
MM161716D	red	0.867	0.872	0.894	0.6	3.1
	green	0.847	0.851	0.901	0.5	6.4
	blue	0.861	0.907	0.806	5.3	6.4
MM161717D	red	0.849	0.853	0.907	0.5	6.8
	green	0.842	0.845	0.877	0.4	4.2
	blue	0.858	0.884	0.832	3.0	3.0
MM162173D	red	0.847	0.845	0.858	0.2	1.3
	green	0.834	0.839	0.824	0.6	1.2
	blue	0.843	0.868	0.804	3.0	4.6
MM162174D	red	0.891	0.876	0.914	1.7	2.6
	green	0.873	0.857	0.898	1.8	2.9
	blue	0.888	0.902	0.827	1.6	6.9
MM162175D	red	0.918	0.904	0.856	1.5	6.8
	green	0.903	0.877	0.884	2.9	2.1
	blue	0.924	0.946	0.777	2.4	15.9
Overall	red				**5.1**	**15.5**
	green				**2.9**	**11.9**
	blue				**6.5**	**19.3**

8.2.4 Center Position Confidence

Table 6. Center position confidence as Δ(Min-Max) the largest distance per measurement in pixel

ImageID	coordinate	phase shift approach	external marker	fast line segments	Δ(Min-Max) [pixel]
MM16708	x	1403.99	1404.04	1404.00	0.13
	y	231.20	231.10	231.08	
MM16709	x	1553.27	1553.27	1553.28	0.10
	y	559.46	559.36	559.38	
MM16710	x	1709.34	1709.15	1709.28	0.27
	y	890.93	890.74	890.82	
MM16711	x	1930.48	1930.43	1930.47	0.11
	y	1170.40	1170.40	1170.30	
MM16712	x	2056.29	2056.29	2056.28	0.20
	y	1376.90	1376.80	1376.70	
MM16713	x	2090.49	2090.54	2090.42	0.16
	y	1572.10	1572.10	1572.00	
MM16714	x	2104.77	2104.91	2104.79	0.24
	y	1839.40	1839.30	1839.20	
MM16715	x	2059.95	2059.98	2060.02	0.21
	y	2123.40	2123.20	2123.20	
MM16716	x	2114.48	2114.49	2114.51	0.20
	y	2449.90	2449.80	2449.70	
MM16717	x	2103.54	2103.66	2103.56	0.23
	y	2800.10	2800.00	2799.90	
MM16718	x	2061.27	2061.20	2061.18	0.31
	y	3148.50	3148.20	3148.20	
MM16173	x	2737.48	2737.40	2737.33	0.16
	y	198.50	198.51	198.45	
MM16174	x	2605.43	2605.33	2605.34	0.17
	y	501.08	500.94	500.95	
MM16175	x	2533.13	2533.13	2533.05	0.14
	y	767.89	767.77	767.79	
MM16176	x	2393.88	2393.83	2393.78	0.22
	y	1056.10	1056.00	1055.90	
MM16177	x	2178.14	2178.12	2178.00	0.33
	y	1311.30	1311.20	1311.00	
MM16178	x	2015.44	2015.42	2015.39	0.11
	y	1579.90	1579.80	1579.80	
MM16179	x	1989.48	1989.35	1989.47	0.24
	y	1857.70	1857.60	1857.50	
MM16180	x	1970.31	1970.31	1970.30	0.30
	y	2128.50	2128.50	2128.20	
MM16181	x	2107.33	2107.44	2107.32	0.32
	y	2377.10	2377.00	2376.80	
Overall	x				**0.21**
	y				

8.2.5 Center Position MTF Distribution

Table 7. Center position MTF distribution as Δ(Min-Max) the largest deviation per MTF measurement in percent

ImageID	channel	phase shift approach	external marker	fast line segments	Δ(Min-Max) [%]
MM16708	red	0.831	0.847	0.841	1.9
MM16708	green	0.816	0.832	0.825	1.9
MM16708	blue	0.833	0.848	0.842	1.8
MM16709	red	0.854	0.863	0.865	1.3
MM16709	green	0.839	0.843	0.845	0.7
MM16709	blue	0.863	0.872	0.876	1.5
MM16710	red	0.922	0.908	0.876	5.0
MM16710	green	0.900	0.882	0.845	6.1
MM16710	blue	0.921	0.906	0.868	5.8
MM16711	red	0.826	0.830	0.830	0.5
MM16711	green	0.822	0.826	0.825	0.5
MM16711	blue	0.833	0.837	0.837	0.5
MM16712	red	0.782	0.788	0.786	0.8
MM16712	green	0.783	0.787	0.786	0.5
MM16712	blue	0.796	0.801	0.799	0.6
MM16713	red	0.727	0.732	0.721	1.5
MM16713	green	0.727	0.732	0.722	1.4
MM16713	blue	0.739	0.745	0.729	2.1
MM16714	red	0.788	0.774	0.769	2.4
MM16714	green	0.796	0.784	0.780	2.0
MM16714	blue	0.810	0.796	0.792	2.2
MM16715	red	0.880	0.831	0.835	5.6
MM16715	green	0.851	0.816	0.819	4.1
MM16715	blue	0.848	0.812	0.816	4.2
MM16716	red	0.888	0.868	0.879	2.3
MM16716	green	0.869	0.849	0.860	2.3
MM16716	blue	0.887	0.869	0.880	2.0
MM16717	red	0.854	0.844	0.846	1.2
MM16717	green	0.844	0.835	0.839	1.1
MM16717	blue	0.863	0.846	0.852	2.0
MM16718	red	0.845	0.807	0.807	4.5
MM16718	green	0.832	0.793	0.793	4.7
MM16718	blue	0.845	0.803	0.803	5.0
MM16173	red	0.853	0.868	0.858	1.7
MM16173	green	0.838	0.847	0.841	1.1
MM16173	blue	0.848	0.865	0.853	2.0
MM16174	red	0.878	0.877	0.876	0.2
MM16174	green	0.859	0.856	0.855	0.5
MM16174	blue	0.874	0.872	0.871	0.3
MM16175	red	0.927	0.921	0.926	0.6
MM16175	green	0.911	0.902	0.910	1.0
MM16175	blue	0.933	0.927	0.932	0.6
MM16176	red	0.808	0.818	0.812	1.2
MM16176	green	0.808	0.817	0.812	1.1
MM16176	blue	0.813	0.822	0.817	1.1
MM16177	red	0.733	0.760	0.738	3.6
MM16177	green	0.747	0.768	0.752	2.7
MM16177	blue	0.761	0.782	0.766	2.7
MM16178	red	0.752	0.740	0.747	1.6
MM16178	green	0.746	0.734	0.742	1.6
MM16178	blue	0.753	0.744	0.750	1.2
MM16179	red	0.774	0.734	0.772	5.2
MM16179	green	0.774	0.735	0.771	5.0
MM16179	blue	0.792	0.750	0.790	5.3
MM16180	red	0.850	0.853	0.834	2.2
MM16180	green	0.843	0.844	0.828	1.9
MM16180	blue	0.856	0.859	0.837	2.6
MM16181	red	0.875	0.851	0.852	2.7
MM16181	green	0.863	0.845	0.844	2.2
MM16181	blue	0.880	0.860	0.858	2.5
Overall	red				**2.3**
	green				**2.1**
	blue				**2.3**

8.2.6 Motion Smear Measurements

Table 8. Quotient M of semi-minor and semi-major ellipse axis for images under kinematic conditions.

Image ID	Quotient of ellipse-axes $M(e_s, e_l)$ with semi-minor axis e_s and semi-major axis e_l
MM00583	0.910
MM00584	0.893
MM00585	0.899
MM00586	0.974
MM00587	0.943
MM00588	0.900
MM00656	0.904
MM00657	0.951
MM00658	0.961
MM00659	0.876
MM00660	0.868
MM02094	0.887
MM02096	0.928
MM02097	0.889
MM02208	0.894
Mean	**0.912**

8.2.7 Exposure Timeline MTF Measurements

Table 9. Exposure timeline MTF measurements as Δ(Mean-Max) the largest deviation per MTF measurement to mean of all four exposure times in percent

ImageID	channel	exp. lowest	exp. medium low	exp. medium hight	exp. highest	Δ(Mean-Max) [%]
MM16152D	red	0.969	0.965	0.966	0.967	0.23
	green	0.922	0.917	0.917	0.916	0.44
	blue	0.956	0.955	0.952	0.953	0.21
MM16605D	red	0.955	0.950	0.948	0.957	0.47
	green	0.915	0.912	0.908	0.917	0.55
	blue	0.950	0.946	0.941	0.951	0.63
MM16701D	red	0.830	0.843	0.837	0.835	0.81
	green	0.792	0.801	0.796	0.795	0.63
	blue	0.819	0.826	0.822	0.820	0.52
MM16829D	red	0.439	0.452	0.468	0.453	3.31
	green	0.442	0.446	0.458	0.447	2.18
	blue	0.439	0.442	0.451	0.442	1.69
MM12553D	red	1.070	1.018	0.978	1.002	5.21
	green	0.977	0.970	0.936	0.951	2.35
	blue	1.035	1.019	0.979	1.009	3.12
MM12924D	red	0.961	0.971	0.966	0.968	0.57
	green	0.924	0.937	0.927	0.924	0.97
	blue	0.972	0.985	0.965	0.967	1.31
MM12043D	red	0.963	0.973	0.966	0.967	0.59
	green	0.922	0.927	0.931	0.918	0.70
	blue	0.958	0.964	0.981	0.956	1.68
P205D	red	0.765	0.705	0.698	0.716	6.10
	green	0.731	0.665	0.663	0.676	6.91
	blue	0.736	0.680	0.677	0.693	5.67
PM205D	red	0.784	0.732	0.709	0.793	6.03
	green	0.711	0.653	0.631	0.720	7.03
	blue	0.732	0.696	0.680	0.724	3.95
RC683D	red	0.961	1.010	1.095	1.021	7.17
	green	0.828	0.831	0.830	0.829	0.18
	blue	0.877	0.930	0.954	0.918	4.65
Overall	red					**3.05**
	green					**2.19**
	blue					**2.34**

8.2.8 MTF Measurements in Relation to Number of Siemens-star Segments

Table 10. MTF Measurements in Relation to Number of Siemens-star Segments

ImageID	channel	16 star segments	36 star segments	Δ(16 segments, 36 segments) [%] (min/max)
MM163212M	red	0.918	0.917	0.1
	green	0.823	0.847	2.8
	blue	0.842	0.849	0.8
MM163212D	red	0.962	0.948	1.5
	green	0.91	0.9	1.1
	blue	0.937	0.937	0.0
MM163152D	red	0.969	0.97	0.1
	green	0.922	0.905	1.8
	blue	0.956	0.953	0.3
MM163159D	red	0.965	0.963	0.2
	green	0.917	0.906	1.2
	blue	0.955	0.95	0.5
MM163206D	red	0.966	0.949	1.8
	green	0.917	0.894	2.5
	blue	0.952	0.939	1.4
MM163717D	red	0.833	0.846	1.5
	green	0.793	0.771	2.8
	blue	0.818	0.828	1.2
MM16583D	red	0.975	0.925	5.1
	green	0.963	0.913	5.2
	blue	0.986	0.947	4.0
MM16585D	red	0.793	0.794	0.1
	green	0.797	0.794	0.4
	blue	0.803	0.799	0.5
MM16587D	red	0.893	0.899	0.7
	green	0.884	0.881	0.3
	blue	0.916	0.897	2.1
MM16659D	red	0.828	0.819	1.1
	green	0.824	0.807	2.1
	blue	0.831	0.844	1.5
MM16095D	red	0.695	0.684	1.6
	green	0.704	0.69	2.0
	blue	0.701	0.691	1.4
MM16308D	red	0.913	0.896	1.9
	green	0.887	0.873	1.6
	blue	0.908	0.887	2.3
MM121638V	red	0.901	0.934	3.5
	green	0.854	0.862	0.9
	blue	0.902	0.937	3.7
MM121951D	red	0.972	1.022	4.9
	green	0.926	0.988	6.3
	blue	0.968	1.055	8.2
P1977L	red	0.61	0.601	1.5
	green	0.758	0.769	1.4
	blue	0.444	0.487	8.8
SG377N	red	1.191	1.143	4.0
	green	1.196	1.146	4.2
	blue	1.164	1.119	3.9
R7673DCB	red	1.013	1.044	3.0
	green	0.956	0.947	0.9
	blue	0.992	0.981	1.1
SA9265D	red	0.387	0.377	2.6
	green	0.358	0.349	2.5
	blue	0.357	0.365	2.2
P1435D	red	0.711	0.721	1.4
	green	0.699	0.713	2.0
	blue	0.738	0.746	1.1
MM165754M	red	0.991	0.948	4.3
	green	0.947	0.926	2.2
	blue	0.965	0.939	2.7
Overall	red			**2.3**
	green			**2.2**
	blue			**2.8**

8.2.9 CTF Measurement of sinusoidal wave and MTF Measurement of square wave Siemens-star

Table 11. CTF measurement of sinusoidal wave and MTF measurement of square wave Siemens-star as Δ(MTF,CTF) in percent

ImageID	channel	MTF square wave	CTF sinusoidal wave	Δ(CTF sinus, MTF square) [%] (min/max)
ZZ661	red	0.900	0.910	1.1
	green	0.874	0.891	1.9
	blue	0.893	0.884	1.0
ZZ858	red	0.809	0.850	4.8
	green	0.808	0.850	4.9
	blue	0.811	0.865	6.2
ZZ094	red	0.903	0.880	2.5
	green	0.890	0.873	1.9
	blue	0.904	0.875	3.2
ZZ095	red	0.689	0.675	2.0
	green	0.694	0.682	1.7
	blue	0.694	0.671	3.3
ZZ094	red	0.913	0.885	3.1
	green	0.896	0.876	2.2
	blue	0.915	0.880	3.8
Overall	red			**2.7**
	green			**2.5**
	blue			**3.5**

8.2.10 Comparison of ground resolved distance for USAF51, Slanted-Edge and Siemens-Star

Table 12. Simultaneous determination of ground resolved distance (GRD) for USAF51, Slanted-Edge and Siemens-Star

ImageID	*USAF chart*	*Slanted-Edge (SLE)*	*Siemens-Star*	*Δ(SLE-Star) [%]*	*Δ(Min-Max)[%]*
Img 1	1.45	1.49	1.52	2.0	4.6
Img 2	1.48	1.51	1.52	0.7	2.6
Img 3	1.55	1.45	1.45	0.0	6.5
Img 4	1.39	1.32	1.34	1.5	5.0
Img 5	1.43	1.38	1.30	5.8	9.1
Img 6	1.45	1.24	1.39	10.8	14.5
Img 7	1.42	1.50	1.53	2.0	7.2
Img 8	1.26	1.15	1.19	3.4	8.7
Img 9	1.24	1.15	1.20	4.2	7.3
Img 10	1.19	1.12	1.15	2.6	5.9
Overall				**3.3**	**7.1**

8.2.11 Comparison of Model-Based MTF and PSF for Slanted-Edge and Siemens-Star

Table 13. Model-PSF compared to measured PSF for Slanted-Edge and Siemens-Star

ImageID	*σ SLE*	*σ Star*	$\Delta(\sigma_m, \sigma_{SLE})$ *[%]*	$\Delta(\sigma_m, \sigma_{Star})$ *[%]*	$\Delta(\sigma_{SLE}, \sigma_{Star})$ *[%]*
0.500	0.609	0.598	17.8	16.4	1.7
0.750	0.894	0.856	16.1	12.4	4.2
1.000	1.093	1.076	8.5	7.1	1.6
1.250	1.301	1.306	3.9	4.3	0.4
1.500	1.546	1.532	3.0	2.1	0.9
1.750	1.739	1.748	0.7	0.1	0.5
Overall					**1.6**

8.2.12 Operator independent measurement repeatability

Table 14. Measurements by different operators, star center by operator (op. X & op. Y), star center by algorithm (alg. X & alg. Y), related MTF10 as mean of all channels, Δmean is mean difference to measurement mean value

ImageID		Op. 1	Op. 2	Op. 3	Op. 4	Op. 5	Op. 6	Op. 7	Op. 8	Op. 9	Op. 10	Δmean
03455	op. X	3868.9	3872.4	3869.0	3869.7	3874.9	3868.4	3867.0	3869.1	3871.2	3869.7	1.8
	op. Y	1033.5	1036.1	1034.4	1033.1	1035.5	1034.7	1032.4	1033.5	1036.4	1032.1	1.2
	alg X	3868.6	3868.7	3868.7	3868.6	3868.6	3868.7	3868.7	3868.6	3868.7	3868.6	0.0
	alg Y	1034.4	1034.5	1034.5	1034.4	1034.4	1034.5	1034.5	1034.4	1034.5	1034.4	0.0
	MTF10	0.764	0.766	0.766	0.764	0.766	0.766	0.764	0.764	0.766	0.764	0.1
P0433	op. X	5204.8	5202.9	5202.9	5205.0	5202.9	5202.9	5201.8	5202.9	5201.9	5200.0	1.1
	op. Y	6282.4	6284.2	6279.4	6278.2	6280.2	6277.4	6280.4	6285.2	6279.4	6279.2	1.9
	alg X	5202.4	5202.4	5202.4	5202.4	5202.4	5202.4	5202.4	5202.4	5202.4	5202.4	0.0
	alg Y	6280.4	6280.4	6280.4	6280.4	6280.4	6280.4	6280.4	6280.4	6280.4	6280.4	0.0
	MTF10	0.775	0.775	0.776	0.775	0.775	0.775	0.776	0.775	0.775	0.776	0.1
00431	op. X	999.2	998.8	996.7	998.9	992.0	999.2	1001.8	997.4	998.9	1002.3	1.9
	op. Y	698.9	700.1	700.1	699.3	700.1	698.9	703.1	701.1	699.3	697.8	1.0
	alg X	999.4	999.4	999.4	999.4	999.4	999.4	999.4	999.4	999.4	999.4	0.0
	alg Y	700.1	700.1	700.0	700.1	700.1	700.0	700.1	700.1	700.0	700.1	0.0
	MTF10	0.823	0.827	0.821	0.824	0.824	0.821	0.823	0.827	0.823	0.827	0.2
01710	op. X	1709.3	1700.6	1713.6	1701.2	1704.3	1704.6	1708.6	1698.2	1700.4	1712.6	5.5
	op. Y	887.8	894.6	892.3	891.4	886.6	892.6	891.3	891.6	894.7	890.3	2.0
	alg X	1709.2	1709.2	1709.2	1709.2	1709.2	1709.2	1709.2	1709.2	1709.2	1709.2	0.0
	alg Y	890.8	890.8	890.8	890.8	890.8	890.8	890.8	890.8	890.8	890.8	0.0
	MTF10	0.893	0.895	0.894	0.895	0.893	0.895	0.894	0.895	0.895	0.894	0.1
01711	op. X	1937.1	1926.3	1931.7	1928.3	1929.2	1937.2	1928.3	1931.8	1928.0	1920.3	3.8
	op. Y	1169.6	1173.9	1170.3	1175.0	1176.1	1167.6	1174.9	1170.0	1176.7	1174.8	3.3
	alg X	1930.5	1930.4	1930.4	1930.4	1930.4	1930.5	1930.4	1930.4	1930.4	1930.4	0.0
	alg Y	1170.4	1170.4	1170.4	1170.4	1170.4	1170.4	1170.4	1170.4	1170.4	1170.4	0.0
	MTF10	0.831	0.828	0.828	0.827	0.828	0.828	0.827	0.831	0.828	0.827	0.1
01717	op. X	2109.4	2104.3	2109.0	2102.5	2108.4	2101.3	2110.1	2102.9	2105.0	2098.5	3.4
	op. Y	2802.9	2795.0	2807.5	2810.1	2803.8	2794.2	2806.5	2808.1	2807.5	2811.9	6.9
	alg X	2103.5	2103.5	2103.5	2103.5	2103.5	2103.5	2103.5	2103.5	2103.5	2103.5	0.0
	alg Y	2799.9	2799.9	2799.9	2799.9	2799.9	2799.9	2799.9	2799.9	2799.9	2799.9	0.0
	MTF10	0.839	0.839	0.841	0.840	0.839	0.840	0.841	0.839	0.839	0.841	0.1
02174	op. X	2605.1	2608.6	2605.3	2607.1	2611.1	2607.6	2607.3	2613.1	2610.1	2608.6	3.1
	op. Y	498.0	513.2	500.0	514.6	505.0	511.2	502.3	512.6	500.9	509.2	6.5
	alg X	2605.3	2605.3	2605.3	2605.3	2605.3	2605.3	2605.3	2605.3	2605.3	2605.3	0.0
	alg Y	501.0	501.0	501.0	501.0	501.0	501.0	501.0	501.0	501.0	501.0	0.0
	MTF10	0.859	0.855	0.861	0.856	0.855	0.861	0.856	0.859	0.855	0.861	0.2
02178	op. X	2014.8	2016.4	2018.6	2020.9	2015.5	2014.6	2016.8	2022.8	2010.6	2019.6	2.9
	op. Y	1585.6	1573.0	1582.7	1581.7	1586.2	1570.3	1587.8	1571.7	1580.6	1579.7	5.0
	alg X	2015.4	2015.4	2015.4	2015.4	2015.4	2015.4	2015.4	2015.4	2015.4	2015.4	0.0
	alg Y	1579.8	1579.8	1579.8	1579.8	1579.8	1579.8	1579.8	1579.8	1579.8	1579.8	0.0
	MTF10	0.739	0.740	0.740	0.739	0.739	0.740	0.740	0.740	0.740	0.739	0.1
02181	op. X	2111.5	2107.6	2111.4	2105.0	2115.1	2106.7	2114.1	2103.3	2101.6	2110.5	3.9
	op. Y	2391.2	2377.0	2375.5	2378.4	2392.1	2370.7	2374.6	2374.8	2376.0	2379.4	4.7
	alg X	2107.3	2107.3	2107.3	2107.3	2107.3	2107.3	2107.3	2107.3	2107.3	2107.3	0.0
	alg Y	2376.8	2376.8	2376.8	2376.8	2376.8	2376.8	2376.8	2376.8	2376.8	2376.8	0.0
	MTF10	0.848	0.850	0.848	0.851	0.848	0.850	0.848	0.851	0.848	0.848	0.1
02173	op. X	2739.2	2742.8	2753.2	2740.1	2731.9	2748.0	2751.4	2740.1	2733.2	2738.4	6.4
	op. Y	194.7	206.0	203.9	194.9	192.7	200.0	202.9	201.8	197.5	195.8	3.9
	alg X	2737.3	2737.3	2737.3	2737.3	2737.3	2737.3	2737.3	2737.3	2737.3	2737.3	0.0
	alg Y	198.5	198.5	198.5	198.5	198.5	198.5	198.5	198.5	198.5	198.5	0.0
	MTF10	0.853	0.852	0.853	0.852	0.853	0.852	0.853	0.853	0.852	0.853	0.1

References

[1] A. Ångström, "On the atmospheric transmission of sun radiation and on dust in the air," *Geografiska Annaler*, vol. 11, no. 2, pp. 156–166, 1929.

[2] D. L. Fried, "Optical resolution through a randomly inhomogeneous medium for very long and very short exposures," *J. Opt. Soc. Am.*, vol. 56, pp. 1372–1379, Oct 1966.

[3] M. Born and E. Wolf, *Principles of Optics*. Cambridge University Press, 1999.

[4] C. S. Williams and O. A. Becklund, *Introduction to the optical transfer function*. Wiley New York etc, 1989.

[5] L. Seidel, "Zur Dioptrik. Ueber die Entwicklung der Glieder 3ter Ordnung, welche den Weg eines ausserhalb der Ebene der Axe gelegenen Lichtstrahles durch ein System brechender Medien bestimmen," *Astronomische Nachrichten*, 1856.

[6] T. Kraft, M. Geßner, H. Meißner, H.-J. Przybilla, and M. Gerke, "Introduction of a photogrammetric camera system for RPAS with highly accurate GNSS/IMU information for standardized workflows," in *EuroCOW 2016, the European Calibration and Orientation Workshop (Volume XL-3/W4)* (J. Skaloud and I. Colomina, eds.), pp. 71–75, Februar 2016.

[7] T. Kraft, M. Geßner, H. Meißner, M. Cramer, H.-J. Przybilla, and M. Gerke, "Evaluation of a metric camera system tailored for high precision UAV applications," in *International Society for Photogrammetry and Remote Sensing*, Volume XLI-B1, 2016, 2016.

[8] H. Jahn and R. Reulke, *Systemtheoretische Grundlagen optoelektronischer Sensoren*. Akademie Verlag, 1995.

[9] H. Meißner, M. Cramer, and B. Piltz, "Benchmarking the optical resolving power of UAV based camera systems," in *UAVg 2017, The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, 42.*, p. 243, Sept. 2017.

[10] E. Abbe, "Beiträge zur Theorie des Mikroskops und der mikroskopischen Wahrnehmung," *Archiv für mikroskopische Anatomie*, vol. 9, no. 1, pp. 413–418, 1873.

[11] L. Rayleigh, "XII. On the manufacture and theory of diffraction-gratings," *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, vol. 47, no. 310, pp. 81–93, 1874.

[12] P. F. Meilan and M. Garavaglia, "Rayleigh criterion of resolution and light sources of different spectral composition," in *Fifth International Topical Meeting on Education and Training in Optics* (C. H. F. Velzel, ed.), vol. 3190, pp. 296 – 303, International Society for Optics and Photonics, SPIE, 1997.

[13] J. Blackledge, *Digital Image Processing: Mathematical and Computational Methods*. Woodhead Publishing Series in Electronic and Optical Materials, Elsevier Science, 2005.

[14] O. Passon, "Gibt es eine Grenze des Auflösungsvermögens?," *IZWT Kolloquium, Wuppertal*, 2016.

[15] A. Glindemann, *Principles of stellar interferometry*. Springer Science & Business Media, 2011.

[16] S. Helmrich, "Improving optical resolution by noise correlation analysis," Master's thesis, University of Heidelberg, 2013.

[17] K. K. Sharma, *Optics: principles and applications*. Elsevier, 2006.

[18] C. M. Sparrow, "On spectroscopic resolving power," *The Astrophysical Journal*, vol. 44, p. 76, 1916.

[19] G. S. Kino and T. R. Corle, *Confocal scanning optical microscopy and related imaging systems*. Academic Press, 1996.

[20] R. Barakat, "Application of apodization to increase two-point resolution by the sparrow criterion. i. coherent illumination," *J. Opt. Soc. Am.*, vol. 52, pp. 276–283, Mar 1962.

[21] W. R. Dawes, "Catalogue of Micrometrical Measurements of Double Stars," vol. 35, p. 137, 1867.

[22] T. S. McKechnie, *General theory of light propagation and imaging through the atmosphere*. Springer, 2016.

[23] J. G. Robertson, "Quantifying resolving power in astronomical spectra," *Publications of the Astronomical Society of Australia*, vol. 30, 2013.

[24] S. Becker, N. Haala, and R. Reulke, "Determination and improvement of spatial resolution for digital aerial images," in *ISPRS Hannover Workshop*, 2005.

[25] E. W. Weisstein, "Gaussian function." http://mathworld.wolfram.com/ GaussianFunction.html, July 2019.

[26] R. Reulke, S. Becker, N. Haala, and U. Tempelmann, "Determination and improvement of spatial resolution of the CCD-line-scanner system ADS40," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 60, no. 2, pp. 81–90, 2006.

[27] G. D. Boreman, *Modulation transfer function in optical and electro-optical systems*, vol. 4. SPIE press Bellingham, WA, 2001.

[28] C. Loebich, D. Wueller, B. Klingen, and A. Jaeger, "Digital camera resolution measurements using sinusoidal siemens stars," in *Digital Photography III*, vol. 6502, p. 65020N, International Society for Optics and Photonics, 2007.

[29] X. Zhang, T. Kashti, D. Kella, T. Frank, D. Shaked, R. Ulichney, M. Fischer, and J. P. Allebach, "Measuring the modulation transfer function of image capture devices: what do the numbers really mean?," in *Image Quality and System Performance IX*, vol. 8293, p. 829307, International Society for Optics and Photonics, 2012.

[30] E. Honkavaara, J. Jaakkola, L. Markelin, and S. Becker, "Evaluation of resolving power and MTF of DMC," *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 36, no. 6, 2006.

[31] C. Goddard Space Flight, *Advanced scanners and imaging systems for earth observations*. NASA SP-335, Washington: Scientific and Technical Information Office, National Aeronautics and Space Administration, 1973.

[32] F. Kharfi, O. Denden, A. Bourenane, T. Bitam, and A. Ali, "Spatial resolution limit study of a ccd camera and scintillator based neutron imaging system according to mtf determination and analysis," *Applied Radiation and Isotopes*, vol. 70, no. 1, pp. 162–166, 2012.

[33] U. Artmann and D. Wueller, "Improving texture loss measurement: spatial frequency response based on a colored target," in *Image Quality and System Performance IX*, vol. 8293, p. 829305, International Society for Optics and Photonics, 2012.

[34] Á. Q. Valenzuela and J. C. G. Reyes, "Basic spatial resolution metrics for satellite imagers," *IEEE Sensors Journal*, vol. 19, no. 13, pp. 4914–4922, 2019.

[35] J. Nakamura, *Image sensors and signal processing for digital still cameras*. CRC press, 2016.

[36] R. Barakat, "Rayleigh wavefront criterion," *Josa*, vol. 55, no. 5, pp. 572–573, 1965.

[37] V. N. Mahajan, "Zernike annular polynomials for imaging systems with annular pupils," *J. Opt. Soc. Am.*, vol. 71, pp. 75–85, Jan 1981.

[38] M. M. Thornton and M. J. Flynn, "Measurement of the spatial resolution of a clinical volumetric computed tomography scanner using a sphere phantom," in *Medical Imaging 2006: Physics of Medical Imaging* (M. J. Flynn and J. Hsieh, eds.), vol. 6142, pp. 707 – 716, International Society for Optics and Photonics, SPIE, 2006.

[39] F. Kharfi, O. Denden, A. Bourenane, T. Bitam, and A. Ali, "Spatial resolution limit study of a ccd camera and scintillator based neutron imaging system according to mtf determination and analysis," *Applied Radiation and Isotopes*, vol. 70, no. 1, pp. 162–166, 2012.

[40] B. Triggs, P. F. McLauchlan, R. I. Hartley, and A. W. Fitzgibbon, "Bundle adjustment—a modern synthesis," in *International workshop on vision algorithms*, pp. 298–372, Springer, 1999.

[41] E. Weitz, "Sift - scale-invariant feature transform homepage," 2019.

[42] H. Moravec, "Towards automatic visual obstacle avoidance," in *Proc. 5th Int. Joint Conf. Art. Intell., 1977*, 1977.

[43] W. Förstner and E. Gülch, "A fast operator for detection and precise location of distinct points, corners and centres of circular features," in *Proc. ISPRS intercommission conference on fast processing of photogrammetric data*, pp. 281–305, Interlaken, 1987.

[44] M. Stephens and C. Harris, "A combined corner and edge detector," in *Alvey vision conference*, vol. 15, 1988.

[45] D. G. Lowe, "Object recognition from local scale-invariant features," in *Proceedings of the seventh IEEE international conference on computer vision*, vol. 2, pp. 1150–1157, Ieee, 1999.

[46] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International journal of computer vision*, vol. 60, no. 2, pp. 91–110, 2004.

[47] I. Rey Otero and M. Delbracio, "Anatomy of the SIFT Method," *Image Processing On Line*, vol. 4, pp. 370–396, 2014.

[48] Z. Zhang, "A flexible new technique for camera calibration," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, pp. 1330–1334, Nov 2000.

[49] M. Lourakis and A. A. Argyros, "Is levenberg-marquardt the most efficient optimization algorithm for implementing bundle adjustment?," in *Tenth IEEE International Conference on Computer Vision (ICCV'05) Volume 1*, vol. 2, pp. 1526–1531, IEEE, 2005.

[50] B. Triggs, A. Zisserman, and R. Szeliski, *Vision Algorithms: Theory and Practice: International Workshop on Vision Algorithms Corfu, Greece, September 21-22, 1999 Proceedings*. Springer, 2003.

[51] N. Qian, "Binocular disparity and the perception of depth," *Neuron*, vol. 18, no. 3, pp. 359–368, 1997.

[52] H. Hirschmüller, M. Buder, and I. Ernst, "Memory efficient semi-global matching," *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 3, pp. 371–376, 2012.

[53] H. Hirschmuller, "Accurate and efficient stereo processing by semi-global matching and mutual information," in *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, vol. 2, pp. 807–814, IEEE, 2005.

[54] R. W. Hamming, "Error detecting and error correcting codes," *The Bell system technical journal*, vol. 29, no. 2, pp. 147–160, 1950.

[55] A. Orych, "Review of methods for determining the spatial resolution of uav sensors," *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. XL-1/W4, 08 2015.

[56] H. Meißner, M. Cramer, and R. Reulke, "Towards standardized evaluation of image quality for airborne camera systems," *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. XLII-1, pp. 295–300, 2018.

[57] USAF, *Military Standard Photographic Lenses*, May 1959. MIL-STD-150A.

[58] N. Koren, "Understanding image sharpness," 2003. http://www.normankoren.com/Tutorials/MTF.html (January 2019).

[59] ISO, "Iso 12233:2017 Photography - Electronic still picture imaging — Resolution and spatial frequency responses," *https://www.iso.org/standard/71696.html*, 2017.

[60] P. D. Burns, "Slanted-edge mtf for digital camera and scanner analysis," in *Is and Ts Pics Conference*, pp. 135–138, SOCIETY FOR IMAGING SCIENCE & TECHNOLOGY, 2000.

[61] S. Birchfield, "Reverse-projection method for measuring camera MTF," *Electronic Imaging*, vol. 2017, no. 12, pp. 105–112, 2017.

[62] T. Choi and D. L. Helder, "Generic sensor modeling for modulation transfer function (MTF) estimation," *South Dakota: Global Priorities in Land Remote Sensing*, 2005.

[63] K. Kohm, "Modulation transfer function measurement method and results for the orbview-3 high resolution imaging satellite," in *Proceedings of ISPRS*, pp. 12–23, 2004.

[64] R. G. Von Gioi, J. Jakubowicz, J.-M. Morel, and G. Randall, "Lsd: a line segment detector," *Image Processing On Line*, vol. 2, pp. 35–55, 2012.

[65] W. S. Russell, "Polynomial interpolation schemes for internal derivative distributions on structured grids," *Applied Numerical Mathematics*, vol. 17, no. 2, pp. 129–171, 1995.

[66] J. W. Coltman, "The specification of imaging properties by response to a sine wave input," *JOSA*, vol. 44, no. 6, pp. 468–471, 1954.

[67] T. Luhmann, "Nahbereichsphotogrammetrie–Grundlagen, Methoden, Beispiele. 4., neu bearbeitete und erweiterte Auflage, Wichmann," 2018.

[68] K. Hirakawa and T. W. Parks, "Adaptive homogeneity-directed demosaicing algorithm," *IEEE Transactions on Image Processing*, vol. 14, no. 3, pp. 360–369, 2005.

[69] E. Chang, S. Cheung, and D. Pan, "Color filter array recovery using a threshold-based variable number of gradients," *Sensors, Cameras, and Applications for Digital Photography*, vol. Proc. SPIE 3650, 1999.

[70] H. S. Malvar, L.-w. He, and R. Cutler, "High-quality linear interpolation for demosaicing of bayer-patterned color images," in *Acoustics, Speech, and Signal Processing, 2004. Proceedings.(ICASSP'04). IEEE International Conference on*, vol. 3, pp. iii–485, IEEE, 2004.

[71] P. Getreuer, "Malvar-he-cutler linear image demosaicking," *Image Processing on Line*, vol. 1, pp. 83–89, 2011.

[72] J. Góźdź, "DCB demosaicing," 2009. https://www.libraw.org/docs/Samples-LibRaw.html.

[73] C. Stamatopoulos, C. Fraser, and S. Cronk, "Accuracy aspects of utilizing raw imagery in photogrammetric measurement," *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 39, p. B5, 2012.

[74] Cmglee, "Comparison of nearest-neighbour, linear, cubic, bilinear and bicubic interpolation methods by CMG Lee.." Wikimedia Common CC BY-SA 4.0, Nov. 2016.

[75] P. Breeuwsma, "Cubic interpolation." https://www.paulinternet.nl/?page=bicubic. Accessed: 2019-09-26.

[76] E. Catmull and R. Rom, "A class of local interpolating splines," in *Computer aided geometric design*, pp. 317–326, Elsevier, 1974.

[77] L. N. Trefethen and D. Bau III, *Numerical linear algebra*, vol. 50. Siam, 1997.

[78] G. C. Birch and J. C. Griffin, "Sinusoidal siemens star spatial frequency response measurement errors due to misidentified target centers," *Optical Engineering*, vol. 54, no. 7, p. 074104, 2015.

[79] H. Meißner, M. Cramer, and R. Reulke, "Towards standardized evaluation of image quality for airborne camera systems," *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. XLII-1, pp. 295–300, 2018.

[80] R. O. Duda and P. E. Hart, "Use of the hough transformation to detect lines and curves in pictures," tech. rep., Sri International Menlo Park Ca Artificial Intelligence Center, 1971.

[81] R. Grompone von Gioi, J. Jakubowicz, J.-M. Morel, and G. Randall, "LSD: a line segment detector," *Image Processing On Line*, vol. 2, pp. 35–55, 2012.

[82] J. B. Burns, A. R. Hanson, and E. M. Riseman, "Extracting straight lines," *IEEE transactions on pattern analysis and machine intelligence*, no. 4, pp. 425–455, 1986.

[83] F. Nex, M. Gerke, F. Remondino, H.-J. Przybilla, M. Bäumker, and A. Zurhorst, "ISPRS benchmark for multi-platform photogrammetry," *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. II-3/W4, pp. 135–142, 2015.

[84] N. Haala and S. Cavegn, "High density aerial image matching: State-of-the-art and future prospects.," *International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences*, vol. 41, 2016.

[85] D. Frommholz, "A synthetic 3d scene for the validation of photogrammetric algorithms," in *ISPRS Geospatial Week 2019*, vol. XLII-2-W13 of *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Copernicus Publications, Juni 2019.

[86] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*. Cambridge University Press, second ed., 2004.

[87] V. Rodehorst, *Photogrammetrische 3D-Rekonstruktion im Nahbereich durch Auto-Kalibrierung mit projektiver Geometrie*. Wissenschaftlicher Verlag Berlin, 2004.

[88] F. Nex and F. Remondino, "UAV for 3D mapping applications: a review," *Applied geomatics*, vol. 6, no. 1, pp. 1–15, 2014.

[89] H.-J. Przybilla, M. Gerke, I. Dikhoff, and Y. Ghassoun, "Investigations on the geometric quality of cameras for UAV applications using the high precision UAV test field zollern colliery," *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. XLII-2/W13, pp. 531–538, 2019.

[90] P. M. Hubel, J. Liu, and R. J. Guttosch, "Spatial frequency response of color image sensors: Bayer color filters and foveon x3," in *Proc. SPIE*, vol. 5301, pp. 402–407, 2004.

[91] M. Gerke and H.-J. Przybilla, "Accuracy analysis of photogrammetric UAV image blocks: Influence of onboard RTK-GNSS and cross flight patterns," *PFG Photogrammetrie, Fernerkundung, Geoinformation*, vol. 2016, pp. 17–30, 03 2016.

[92] M. Cramer, H.-J. Przybilla, and A. Zurhorst, "UAV Cameras: Overview and geometric calibration benchmark." submitted to this UAV-g 2017 conference, Bonn, Germany, 2017.

[93] E. Rupnik, F. Nex, I. Toschi, and F. Remondino, "Aerial multi-camera systems: Accuracy and block triangulation issues," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 101, pp. 233–246, 2015.

[94] M. Li and T. Q. Nguyen, "Markov random field model-based edge-directed image interpolation," *IEEE Transactions on Image Processing*, vol. 17, no. 7, pp. 1121–1128, 2008.

[95] Q. Wang and R. Ward, "A new edge-directed image expansion scheme," in *Proceedings 2001 International Conference on Image Processing (Cat. No.01CH37205)*, vol. 3, pp. 899–902 vol.3, 2001.

[96] J.-G. Leu, "Edge sharpening through ramp width reduction," *Image and Vision Computing*, vol. 18, no. 6, pp. 501 – 514, 2000.

[97] PhaseOne, "Capture One." https://www.captureone.com/en/, Nov. 2019.

[98] DJI, "Dji drone series." https://www.dji.com/en/, Nov. 2019.

[99] X. Lu, H. Yuan, P. Yan, Y. Yuan, and X. Li, "Geometry constrained sparse coding for single image super-resolution," in *2012 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1648–1655, 2012.

[100] L. Zhang and X. Wu, "An edge-guided image interpolation algorithm via directional filtering and data fusion," *IEEE transactions on Image Processing*, vol. 15, no. 8, pp. 2226–2238, 2006.

[101] C. Dong, C. C. Loy, K. He, and X. Tang, "Learning a deep convolutional network for image super-resolution," in *European conference on computer vision*, pp. 184–199, Springer, 2014.

[102] J. Kim, J. Kwon Lee, and K. Mu Lee, "Accurate image super-resolution using very deep convolutional networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 1646–1654, 2016.

[103] B. Lim, S. Son, H. Kim, S. Nah, and K. Mu Lee, "Enhanced deep residual networks for single image super-resolution," in *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*, pp. 136–144, 2017.

[104] B. K. Horn and B. G. Schunck, "Determining optical flow," in *Techniques and Applications of Image Understanding*, vol. 281, pp. 319–331, International Society for Optics and Photonics, 1981.

[105] D. Fleet and Y. Weiss, "Optical flow estimation," in *Handbook of mathematical models in computer vision*, pp. 237–257, Springer, 2006.

[106] S. S. Beauchemin and J. L. Barron, "The computation of optical flow," *ACM computing surveys (CSUR)*, vol. 27, no. 3, pp. 433–466, 1995.

[107] B. D. Lucas, T. Kanade, *et al.*, "An iterative image registration technique with an application to stereo vision," 1981.

[108] B. Horn, B. Klaus, and P. Horn, *Robot vision*. MIT press, 1986.

[109] D. Shepard, "A two-dimensional interpolation function for irregularly-spaced data," in *Proceedings of the 1968 23rd ACM national conference*, pp. 517–524, 1968.

[110] S. Farsiu, M. Elad, and P. Milanfar, "Multiframe demosaicing and super-resolution of color images," *IEEE transactions on image processing*, vol. 15, no. 1, pp. 141–159, 2005.

[111] S. Farsiu, M. D. Robinson, M. Elad, and P. Milanfar, "Fast and robust multiframe super resolution," *IEEE transactions on image processing*, vol. 13, no. 10, pp. 1327–1344, 2004.

[112] B. Goldlücke, M. Aubry, K. Kolev, and D. Cremers, "A super-resolution framework for high-accuracy multiview reconstruction," *International journal of computer vision*, vol. 106, no. 2, pp. 172–191, 2014.

[113] M. J. Black and P. Anandan, "A framework for the robust estimation of optical flow," *1993 (4th) International Conference on Computer Vision*, pp. 231–236, 1993.

[114] T. Brox, A. Bruhn, N. Papenberg, and J. Weickert, "High accuracy optical flow estimation based on a theory for warping," in *European conference on computer vision*, pp. 25–36, Springer, 2004.

[115] M. Cramer and B. Leinss, "Welche ist am besten?–Anmerkungen zur Auswahl von Kamerasystemen in der UAS-Luftbildphotogrammetrie," *UAV-Vermessung mit unbemannten Flugsystemen: Beiträge zum*, vol. 148, 2016.

[116] T. Luhmann, C. Fraser, and H.-G. Maas, "Sensor modelling and camera calibration for close-range photogrammetry," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 115, pp. 37–46, 2016.

[117] M. K. Transtrum and J. P. Sethna, "Improvements to the Levenberg-Marquardt algorithm for nonlinear least-squares minimization," *arXiv preprint arXiv:1201.5885*, 2012.

[118] M. Rothermel, *Development of a SGM-based multi-view reconstruction framework for aerial imagery*. PhD thesis, University of Stuttgart, 11 2016.

[119] H. Hirschmüller, "Stereo processing by semiglobal matching and mutual information," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 30, pp. 328–341, feb 2008.

[120] H. Hasegawa, K. Matsuo, M. Koarai, N. Watanabe, H. Masaharu, and Y. Fukushima, "Dem accuracy and the base to height (b/h) ratio of stereo images," *International Archives of Photogrammetry and Remote Sensing*, vol. 33, no. B4/1; PART 4, pp. 356–359, 2000.

[121] T. Luhmann, *Erweiterte Verfahren zur geometrischen Kamerakalibrierung in der Nahbereichsphotogrammetrie*. Verlag der Bayerischen Akademie der Wissenschaften in Kommission beim Verlag ..., 2010.

[122] W. Boehler, M. B. Vicent, A. Marbs, *et al.*, "Investigating laser scanner accuracy," *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 34, no. Part 5, pp. 696–701, 2003.

Selbständigkeitserklärung

Ich erkläre hiermit, dass ich die vorliegende Arbeit selbständig verfasst und noch nicht für andere Prüfungen eingereicht habe. Sämtliche Quellen einschließlich Internetquellen, die unverändert oder abgewandelt wiedergegeben werden, insbesondere Quellen für Texte, Grafiken, Tabellen und Bilder, sind als solche kenntlich gemacht. Mir ist bekannt, dass bei Verstößen gegen diese Grundsätze ein Verfahren wegen Täuschungsversuchs bzw. Täuschung eingeleitet wird.

Berlin, den February 18, 2021 .....